Mechanical Engineering

Transactions of the ASME

TRANSACTIONS OF THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS

VOLUME 80 — FEBRUARY 1958 — NUMBER 2

Transactions

of The American Society of Mechanical Engineers

Published on the tenth of every month, except March, June, September, and December

Published monthly by The American Society of Mechanical Engineers. Publication office at 20th and Northampton Streets, Easton, Pa. The editorial department is located at the headquarters of the Society, 29 West Thirty-Ninth Street, New York 18, N. Y. Cable address, "Mechanaut," New York. Price $1.50 a copy, $12.00 annually for Transactions and the *Journal of Applied Mechanics*, to members, $1.00 a copy, $6.00 annually. Add $1.50 for postage to all countries outside the United States, Canada, and Pan American Union. Changes of address must be received at Society headquarters seven weeks before they are to be effective on the mailing list. Please send old as well as new address. . . . By-Law: The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B13, Par. 4). . . . Entered as second-class matter March 2, 1928, at the Post Office at Easton, Pa., under the Act of August 24, 1912. . .

Designing Thermocouples for Response Rate

By R. J. MOFFAT,[1] WARREN, MICH.

Accurate information about thermocouple response rate is of value both to the designer of engine controls and to the test engineer. The response performance of a probe can be described in terms of its "characteristic time," which can be determined experimentally. The characteristic time of a probe is not a constant, but is affected by the mass velocity and temperature of the gas stream in which it is used. Analysis of experimental data resulted in an empirical equation for characteristic time in terms of the geometry and flow conditions. This equation predicts the characteristic time within 10 per cent for bare wire-loop-junction thermocouples from 0.016 to 0.051-in-diam wire, mass velocities from 3 to 50 lb/sec ft² and temperature from 160 to 1600 F. All data were taken at a static pressure of 1 atm. The effects of manufacturing tolerances and engine environmental conditions also have been investigated. Data are presented concerning the effect of junction-weld bead size, junction exposed length, junction orientation, and radiative heat transfer from the junction.

Definition of Characteristic Time

CONSIDERABLE effort has been put into the study of the transient response of thermocouples. One reason for this is the growing interest in temperature-sensitive controls for jet engines. The advantage of a temperature-sensitive control is obvious—it controls by the variable which requires control. The disadvantages are chiefly in the sensing element. An ideal sensing element would be instantly aware of any change in gas temperature, and would follow accurately the temperature no matter how rapidly it changed. Unfortunately no such ideal sensing element is available. Anything which has mass requires a finite time to change its temperature, the length of time depending on its heat capacity and on how fast heat is being added to it. In terms of a thermocouple, or any other immersion element, this means that if the temperature of the gas is changing the thermocouple will "lag" and not follow the change exactly. This lag is important in control work since the control is not aware of a change in gas temperature until the signal from the sensing element changes. The lag is also important in analyzing transient temperature records made on a test engine. The recorded trace represents thermocouple temperature, not gas temperature. Owing to the lag there may be considerable difference between gas temperature and thermocouple temperature.

A temperature-time record made during the starting cycle of a large jet engine is represented in Fig. 1. Two thermocouple traces were recorded, from thermocouples of different sizes and in different locations. The ordinate is temperature in degrees F, the abscissa, time in seconds after ignition. The peak indicated temperature was just over 2000 F. The peak corrected temperature was over 3000 F. The only point on this record where the indicated traces actually mean gas temperature is where the traces are horizontal—where the thermocouple temperature is not changing with time. The two thermocouples gave different traces, even though exposed to the same gas stream, due to their different response rates. Thus, before a transient temperature record can yield information about gas temperature it must be corrected for the lag of the thermocouples. The data for this figure were taken with bare-wire thermocouples with relatively slight lag, 20 or 40 deg for a 10 deg per sec rise in thermocouple temperature. For some probes now in use the lag would be considerably greater. Transient records from such probes seriously distort the true gas-temperature history.

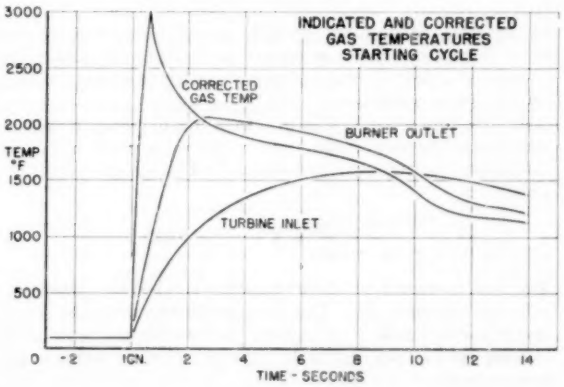


Fig. 1

The transient behavior of a thermocouple can be defined by specifying its "characteristic time," usually represented by the Greek letter τ (tau). This parameter has two useful definitions which may be developed from a basic heat balance.

Considering heat transfer by convection only (no radiation, no conduction) the rate of heat addition equals the rate at which heat is being stored in the thermocouple junction

$$h_c A_e (T_G - T_J) = Mc \frac{dT_J}{dt} \quad \ldots\ldots\ldots\ldots [1]$$

where

h_c = coefficient of heat transfer by convection, Btu/sec sq ft deg R
A_e = heat-transfer area of junction, sq ft
T_G = gas temperature, deg R
T_J = junction temperature, deg R
M = mass of junction, lb
c = specific heat of junction, Btu/lb deg R

It may appear that M and A_e are not defined for a thermocouple junction, but they appear only in a ratio M/A_e, which is easily found. Defining, arbitrarily for the moment, the collection of terms $Mc/h_c A_e$ as τ and rearranging

$$T_G = T_J + \tau \frac{dT_J}{dt} \quad \ldots\ldots\ldots\ldots [2]$$

This equation provides a very useful definition of τ; the number of degrees of lag, per-degree-per-second temperature change. It is by use of this relationship that raw data from an engine transient

[1] Senior Research Engineer, Gas Turbines Department, Research Staff, General Motors Corporation, GM Technical Center.

Contributed by the Gas Turbine Power Division and presented at the Gas Turbine Power Conference, Detroit, Mich., March 18–21, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 22, 1957. Paper No. 57—GTP-8.

can be corrected to yield the gas-temperature history. Given a temperature-time record, the true temperature at any instant can be found by adding to the indicated value, point by point, a correction equal to τ times the indicated rate of change of temperature. This, of course, assumes that there are no other corrections required, such as for radiation, conduction, or velocity effects.

To return to the general Equation [1], solving it for T_J for the case of a step change in gas temperature from T_{G0} to T_G. Initially the thermocouple will indicate T_{J0} which will be assumed equal to T_{G0} (no radiation, conduction, or velocity correction). The indicated temperature at any time t, after the step, is given by

$$T_J = T_{J0} + (T_G - T_{G0})\left(1 - e^{-t/\frac{Mc}{h_c A_c}}\right) \quad [3]$$

$$\frac{T_J - T_{J0}}{T_G - T_{G0}} = \left(1 - e^{-t/\frac{Mc}{h_c A_c}}\right) \quad [4]$$

In Equation [4] the terms Mc/h_cA_c appear in the exponent of e in the same arrangement as implied in Equation [2], and in the position occupied by "characteristic time" in the similar equation for a charging condenser.

As t, the time after the step, goes from zero to infinity it must at some instant be numerically equal to the value of Mc/h_cA_c. At this instant the exponent of e becomes (-1) and Equation [4] becomes

$$\frac{T_J - T_{J0}}{T_G - T_{G0}} = 1 - \frac{1}{e} = 0.632 \quad [5]$$

The time, in seconds, at which this occurs is called the "characteristic time" of the probe. This provides the most common definition of characteristic time; namely, the time required for a probe to complete 63.2 per cent of its response to a step change in gas temperature.

This definition is most useful in comparing probes for transient use: low characteristic time means rapid response. It is also the basis for experimental determination of τ.

If the characteristic time τ is known, and "perfect," the response of a probe can be found from Equation [2] for any definable function of gas temperature, and gas temperature at any instant can be calculated from the behavior of the thermocouple temperature. However, there are several conditions which must be met before τ for any practical probe is perfect enough for such application:

1 The temperature-emf characteristics of the thermocouple must be linear.

2 The response of the thermocouple to changes in gas temperature must fit a first-order differential equation, Equation [2].

3 The temperature must be uniform across the wire at the junction.

4 There must be no heat transfer to the junction by radiation or conduction.

There is an additional complication; the characteristic time of a probe is not a constant. It is affected not only by the physical size and shape of the probe, but also by the flow conditions in which the probe is used. The variation in τ is chiefly due to variation in h_c, the convective heat-transfer coefficient. Any flow condition which causes h_c to change will cause a change in τ. Two parameters appear to be sufficient to describe the flow; namely total temperature, and mass velocity. Total temperature is the sum of the static temperature of the gas and the temperature equivalent of its kinetic energy. Mass velocity is a flow-concentration parameter, the pounds of gas flow per second per square foot of flow area.

Effect of Total Temperature, Mass Velocity, and Wire Diameter

The effect of total temperature and mass velocity on the characteristic time of a 16-gage bare-wire loop junction is shown in Fig. 2. The data cover temperatures from 160 to 1600 F and mass velocities from 2 to 50 lb/sec sq ft. Characteristic time for this probe can have any value from 0.62 to 5.5 sec, depending on the mass velocity and temperature. Characteristic time decreases as the temperature goes up, and also as the mass velocity goes up.

Diameter also affects the characteristic time, as might be expected. The variation is shown in Fig. 3, representing data taken at 800 F with bare-wire-loop junctions from 16 to 26 gage (0.051–0.016 in. diam). Each wire diameter has been tested at 160, 800, 1200, and 1600 F at the various mass velocities. The characteristic time was determined by direct measurement from the emf-time record resulting from a step change, according to Equation [5]. The magnitude of the step change was large—from an initial temperature of 200–500 F for the high-temperature runs. This may have obscured somewhat the effect of temperature level. Until such time as data are available for small step changes, the following empirical equation is proposed for bare-wire loop junctions under ideal conditions

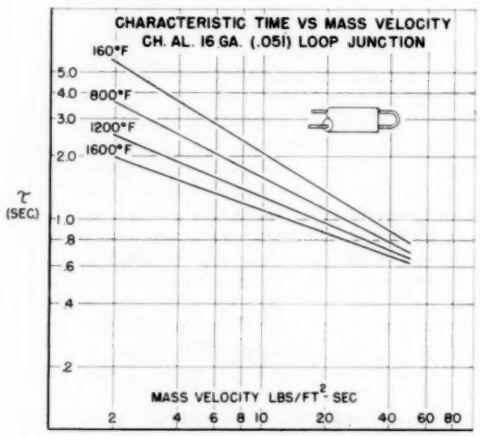


Fig. 2

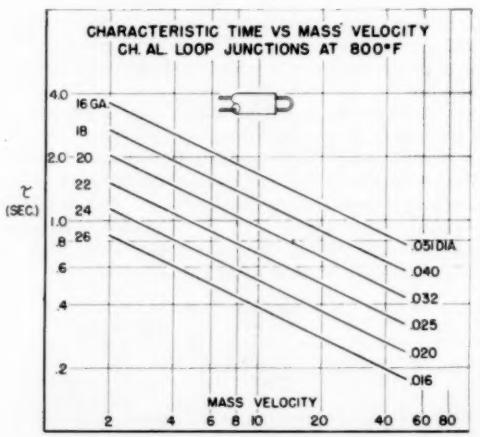


Fig. 3

$$\tau = \frac{3.5 \times 10^5 \rho c d^{1.25}}{T} G^{-15.8/\sqrt{T}} \quad [6]$$

where

τ = characteristic time, sec
d = wire diameter, in.
G = mass velocity, lb/sec sq ft
T = total temperature, deg R
ρ = average density for the two wires, pcf
c = average specific heat for the two wires, Btu/lb deg F

For the case of chromel-alumel, where ρ = 540 lb and c = 0.116 Btu/lb deg F, this becomes

$$\tau = \frac{2.19 \times 10^5 d^{1.25}}{T} G^{-15.8/\sqrt{T}} \quad [7]$$

Fig. 4 compares test data for a 16-gage loop junction with values calculated by this equation. In this figure only, calculated values are shown as solid lines, test data as individual points. No calculated values are shown in any other figure except Fig. 1.

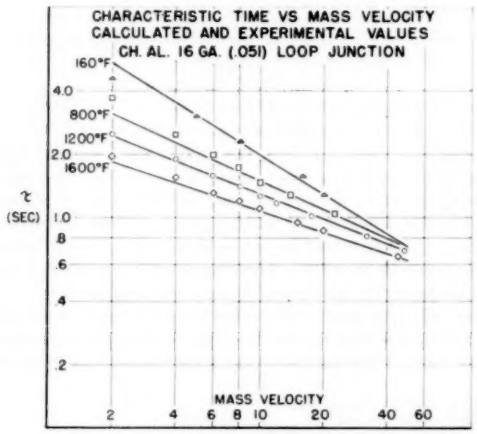


FIG. 4

Examination of Equation [7] reveals several trends of characteristic-time variation within the range of this equation. τ is directly proportional to the average density and specific heat of the material forming the junction. It is proportional to the diameter of the wire to the 1.25 power for any condition of temperature and mass velocity. τ decreases as the temperature rises—but not directly, since temperature occurs also in the exponent of mass velocity. The variation of τ with G becomes less as the temperature level rises, as demonstrated by the lower slope of the 1600-deg F line in Fig. 2.

It must be emphasized that this equation is empirical, and applies only for the ideal case; bare-wire loop junctions, of uniform wire diameter, with no conduction and no radiation effect, in a free stream of uniform conditions.

Deviations Causing Errors

It is not always possible, or economical, to attain ideal conditions in an engine application. Any deviation from ideal conditions which violates one of the initial assumptions may result in characteristic time different from that predicted by Equation [6]. To date, five common deviations have been investigated. Of these, three have been found to be significant (conduction, weld-bead size, and junction shape) and two have little effect (radiation and orientation of the junction in the gas stream).

Conduction. Conduction is the most troublesome factor. Although a given probe may have little or no conduction error at steady state, it may be affected greatly during transients if the exposed wire is not sufficiently long. Conduction from the loop to the stem causes an increase in characteristic time and a loss of first-order characteristic. The effect is most pronounced at low mass velocities, and causes a distortion of the τ versus G curve, raising it at the low G end.

This effect is due to conduction of heat along the wire of the loop to the stem of the probe. Because of this conduction, the junction temperature is partly determined by the temperature of the stem. Since the stem is relatively heavy, it responds slowly to a change in gas temperature and tends to hold back the junction.

The conduction effect may be eliminated by exposing more wire to the gas stream. The length required is a function of wire diameter and the local mass velocity. At a mass velocity of 5 lb/sec per sq ft, a loop 5 wire diameters long will be within the tolerance allowed by Equation [6] for chromel-alumel. The effect of short loop lengths on characteristic time is shown in Fig. 5 for

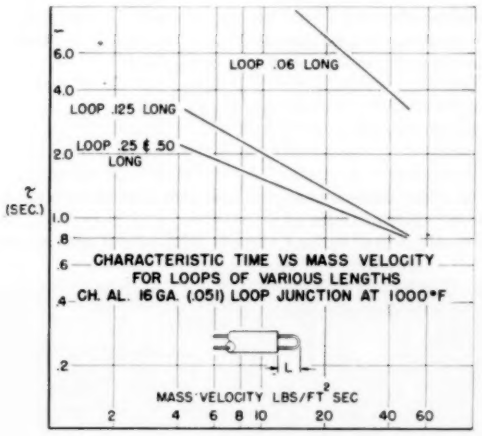


FIG. 5

16-gage bare-wire loop junctions of the chromel-alumel. Materials of higher thermal conductivity would require proportionately longer loops.

Weld-Bead Size. The second significant factor was found to be weld-bead size. Equation [6] was based on experimental probes, with uniform wire diameter throughout the loop. In production thermocouples, some weld bead must be tolerated at the junction. It is to be expected that this extra mass of metal would raise the characteristic time of the junction.

Fig. 6 shows the effect of weld-bead size on the characteristic time of an 18-gage bare-wire loop junction. Although an increase in weld-bead size does raise the characteristic time, the

TABLE 1 CHARACTERISTIC TIMES AT G = 5, T = 160

Diameter	τ, sec For 0.040-wire-diam weld-bead size, as given in col. 1	τ, sec For variable wire diam, as given in col. 1, no weld bead
0.040.....	2.4	2.4
0.048.....	2.6	3.0
0.060.....	2.8	3.9

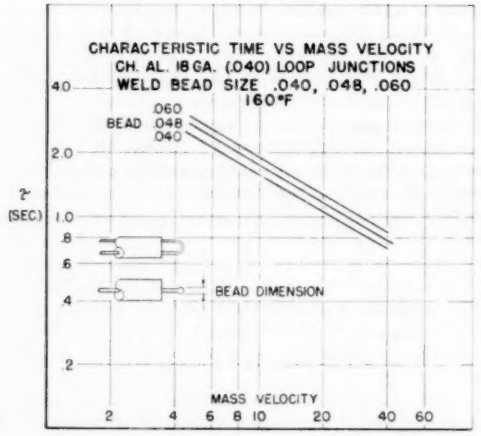


FIG. 6

effect is by no means as significant as would be a corresponding increase in wire diameter. For instance, Table 1 gives the values of τ for three weld-bead sizes on 0.040-diam wire compared with τ for uniform diameter wire of the same diameter as the weld beads.

For the range of sizes given in Table 1 the effect of the weld bead can be approximated by the following equation

$$\frac{\tau}{\tau_0} = \left(\frac{D}{d}\right)^{0.375} \quad \text{..... [8]}$$

where

τ = characteristic time, sec, with a bead of diameter D and a wire of diameter d
τ_0 = characteristic time, seconds, with wire of uniform diameter d, no weld bead
D = weld-bead diameter, in.
d = wire diameter, in.

It is strongly recommended that Equation [8] be used only to set limits on D/d for a given tolerance on τ/τ_0. For instance, to limit the increase in τ to 10 per cent due to the weld bead, D/d must be kept less than 1.29. For 0.040-in-diam wire, then, the bead must not exceed 0.051 in. Weld beads do not necessarily have well-defined shapes, hence D can only be estimated, and D/d should be taken as an upper limit rather than a design value.

Shape of Junction. Independent of junction length or weld bead size, the shape of the junction itself can affect characteristic time. The best example is the "twisted junction," where the wires are tightly twisted together for two or more turns. For such a junction, the characteristic time must be calculated using an "effective wire diameter" of $1.5d$ in Equation [6]. Thus, for a twisted junction made of 0.040-in. wire, τ would be calculated using $d = 0.060$ in. The increase in characteristic time as a result of twisting the wires does not appear to extend to wires which are merely close together. Tests conducted on junctions where the two sides of the loop were parallel and one wire diameter apart showed no increase in characteristic time as compared to open loops of generous radius. To date, no junction shape made with round wire has significantly deviated from Equation [6] if the wires were sufficiently long, and at least one diameter apart. The same applies to double-junction probes; the proximity of the two junctions does not appear to affect their performance, so long as they are at least one wire diameter apart and meet the other requirements of loop length and weld-bead size.

Fig. 7 illustrates several shapes which have been tested, showing their comparative values of characteristic time.

Radiation. Radiation from the probe to the walls is another factor which could affect characteristic time, and thus change the performance of a probe. A series of tests was conducted to evaluate this effect. A bare 16-gage round-wire loop-junction thermocouple was tested for response rate with different rates of radiation loss to the walls. The data are summarized in Table 2.

Although the radiation did have a measurable effect, it was small at this temperature level. In the most severe case there was only a 10 per cent increase in τ. This should not be interpreted to mean that there was little or no radiation error in the thermo-

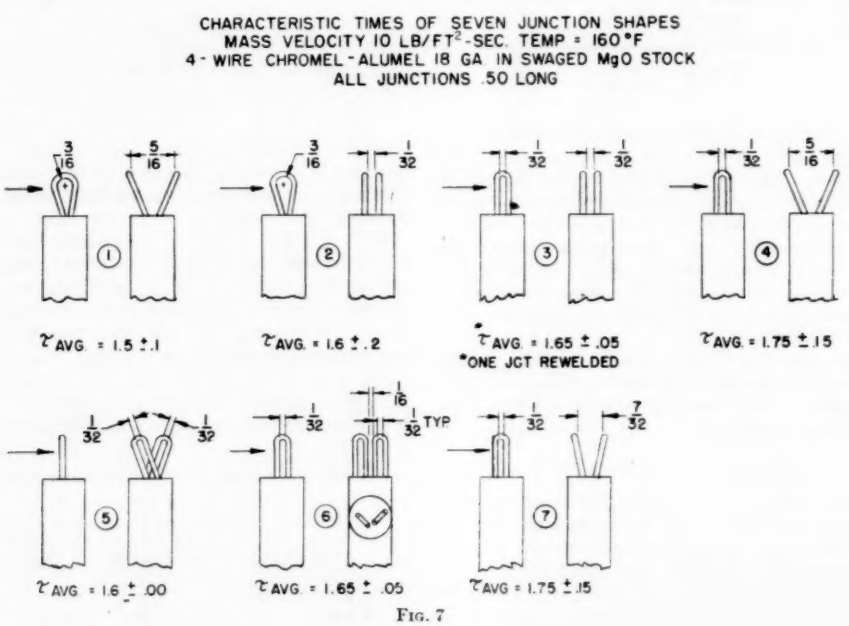


FIG. 7

TABLE 2 EFFECT OF RADIATION TO COLD WALLS ON CHARACTERISTIC TIME; 1600 F GAS TEMPERATURE

G	T_{wall}, deg F	τ, sec
5	635	1.62
5	1222	1.50
10	810	1.28
10	1383	1.16
20	950	0.92
20	1460	0.89

couple signal. It means that the thermocouple signal could be corrected for response rate by using the same value of τ as in the absence of radiation. The temperature resulting from this correction would not be the true gas temperature. It would be the temperature the thermocouple would have indicated if it had been able to respond instantaneously, and thus would include radiation error.

Orientation. One of the uncontrollable conditions in an engine application is the flow angle of the gas stream with respect to the thermocouple. Since this angle may change with engine operating condition, it seemed advisable to check its effect on characteristic time. Tests were conducted on a bare-wire loop junction at 1000 F, with results as indicated in Table 3.

TABLE 3 EFFECT OF JUNCTION ORIENTATION ON CHARACTERISTIC TIME

Plane of loop	Mass velocity, lb/sq ft-sec	τ, sec
Parallel to flow	5	2.39
45 deg to flow	5	2.31
90 deg to flow	5	2.37
Parallel to flow	10	1.80
45 deg to flow	10	1.77
90 deg to flow	10	1.77

From the data in Table 3, it appears that the effect of junction orientation is negligible, at least for relatively open junctions. The junction used was 13 wire diameters long, with the wires parallel, and two wire diameters apart. More dense shapes may be affected by orientation to a greater extent, as might the shorter junctions.

SUMMARY

In summary then, the response rate of a bare-wire loop-junction thermocouple may be predicted by the following equation

$$\tau = \frac{3.5 \times 10^3 \rho c d^{1.25}}{T} G^{-15.8/\sqrt{T}} \quad [9]$$

where

τ = characteristic time, sec
ρ = average density of materials, pcf
c = average specific heat, Btu/lb deg F
d = wire diameter, in.
G = mass velocity, lb/sq ft sec
T = total temperature, deg R

When certain conditions are met, the equation is accurate to within 10 per cent over the following range:

Temperature: From 160 to 1600 F
Wire diameter: From 0.016 to 0.051 in.
Mass velocity: From 3 to 50 lb/sq ft sec
Pressure: 1 atm, static

The necessary conditions are as follows:

1 The junction must be sufficiently long to eliminate conduction effects. The required length is a function of wire diameter and mass velocity. At a mass velocity of 5 lb/sq ft sec, a length equal to 5 wire diameters is required for chromel-alumel. Lower mass velocities require longer loops, as do materials of higher thermal conductivity.

2 Wire diameter must be uniform in the region of the junction. If the junction has an appreciable weld bead, the characteristic time will be increased. The increase is approximately given by

$$\frac{\tau}{\tau_0} = \left(\frac{D}{d}\right)^{0.375} \quad [10]$$

where

τ = characteristic time, seconds, with weld bead
τ_0 = characteristic time, seconds, no weld bead
D = diameter of weld bead, in.
d = wire diameter, in.

3 The junction must be a bare-wire loop junction, or similarly open shape. If the wires of the junction are less than one wire diameter apart, the characteristic time may be increased. Twisting the junction increases the effective diameter of the wire to $1.5d$ for use in Equation [9].

Two environmental factors which are of secondary importance to the response of a thermcouple and which may be neglected are radiation to the walls, and the orientation of the junction with respect to the flow.

TEST EQUIPMENT AND METHOD

Data presented in this paper were taken with the apparatus shown in Fig. 8. This consists of an insulated and electrically heated reference section, at right, and the response-rate test section, at left. The reference section provides an error-free environment for measuring gas temperature, owing to its large diameter and electrically heated walls. The test section consists of a 4-in. gas duct, response-rate sheath-actuator, and electrically heated insulating covers (one of which was removed for the photograph).

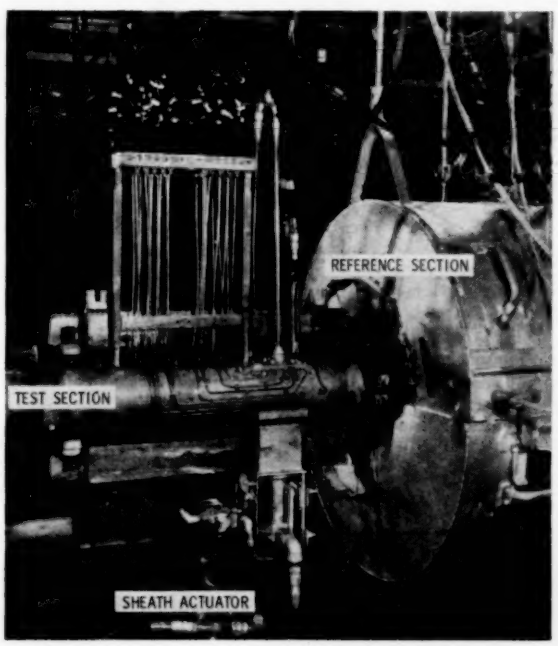


FIG. 8 APPARATUS USED IN TESTS

A convergent nozzle of rectangular section (1.375 × 3.5 in.) is located in the inlet end of the test section. The test thermocouple is located 1 in. downstream from the exit plane of the nozzle. Test-section wall temperatures are measured for 80 per cent of the field of view of the test thermocouple. The response-rate cooling sheath may be raised to cover the test probe and can be retracted in 0.015 sec by the air cylinder. Cooling is achieved by flowing air through the sheath when it is in the raised position.

When the sheath is suddenly retracted, the probe is subjected to a step change in gas temperature, at the flow conditions established in the test section. The output of the thermocouple is amplified by a breaker type d-c amplifier and recorded on a direct-writing oscillograph. The characteristic time is measured directly from the emf versus time record.

All data were taken at static pressures of 1 atm in the test section.

Discussion

G. E. Glawe.[2] The paper is of general interest because it includes information which is useful for design, construction, and application considerations.

A few specific comments follow:

It is thought that after Equation [6] some general discussion of its implications should be noted. For instance, the time constant can be reduced for a given configuration by choosing thermocouple material whose product of ρC is lower (such as platinum rhodium-platinum) or by using a small-diameter wire. The latter case, of course, must be a compromise with aerodynamic loading, life expectancy at operating temperature, and so on.

From Fig. 5, the immersion length L is taken as the projected linear distance from the end of the support to the junction. In a conduction analysis the immersion length to consider is the length from the end of the support to the junction. The L of Fig. 5 is consistent only if the end of the loop has a small radius relative to wire diameter; which does happen to be the case for the majority of the tests reported.

A question, which sometimes arises in regard to the L-dimension, is whether there is flow interference from the support for short-wire immersion lengths. This might lead to an investigation of optimum L to support diameter ratio.

Referring to "effect of radiation on characteristic time," it can be shown that under the conditions of convective heat transfer and radiation

$$\tau = \frac{(d/4)\rho C}{h_c + 4\sigma\epsilon T_j^3}$$

For bare wires in transverse flow, h_c can be evaluated from the Nusselt-Reynolds number relation such as given in NACA TN 2599. This relation is $hD/k = 0.43\ (GD/\mu)^{1/2}$. This theoretical relation predicts the same order of magnitude in the variation of time constant with variation in radiation as shown by Table 2 of the paper. This effect, of course, will become more significant at higher temperatures. Also, a T_j column might be included to advantage in Table 2 to show the actual test-junction temperature variation with variation in mass flow and wall-temperature depression.

J. H. Weaving.[3] The writer finds the paper very interesting and there is no doubt of the importance of being able to follow the actual temperature under transient conditions in the starting of gas turbines and similar fluctuating circumstances. The writer feels, too, that the data produced are accurate and would form a most useful background for such calculations. However, he considers that one is left slightly in the air as the data are not applied to an actual case, unless Fig. 1 is such an application; if so, no reference is made to the method as to how the actual gas temperature is calculated. Figs. 2–6 give basic data on the base of mass velocity, but to apply these data to an actual case we still must solve Equation [2] where τ is given by Equation [9] which contains G and T, both of which are variables. This means, first, that Equation [2] cannot be solved by analytical means and also a knowledge of G, the mass velocity, with respect to time is required. The problem of obtaining the true temperature is thus still a formidable one. Presumably when G, as a function of time, has been obtained experimentally or estimated by calculation, Equation [2] may be integrated graphically to give the true temperature as a function of time.

Author's Closure

There is some question in my mind as to the existence of an optimum ratio of L to support diameter for the loop length effect discussed in this paper, which is primarily due to the conduction of heat from the junction to the support. If the support is sufficiently large to act as a heat sink for this process, then a further increase in support size would not change the effect, although it would change the ratio of L to support diameter. I feel that Mr. Glawe's comment is more applicable at high Mach number flows where there might be local shock interference due to the support.

With reference to Mr. Weaving's comment, it is certainly true that the determination of true temperature from a transient trace is still difficult. The mass velocity must be known, as well as the indicated temperature and rate of change of indicated temperature. The problem is much as Mr. Weaver has stated. The true temperature must be obtained by successive approximation, since τ will change with each new estimate of temperature.

Admittedly, there is a good deal of material which could have been included in this paper—the technique of applying τ to gas temperature determination, the use of τ to calculate radiation error (by determining h_c), the use of τ to determine basic heat transfer relationships such as Nusselt number, and so on. The field is so broad, however, that the result would be unwieldy. Consequently, this paper was trimmed of all material not directly applicable to "Designing Thermocouples for Response Rate." I hope that future papers may expand upon the many applications of τ data which can be made.

[2] Research Engineer, National Advisory Committee for Aeronautics, Cleveland, Ohio.

[3] The Austin Motor Company, Ltd., Longbridge, Birmingham, England.

Analysis of Incompressible, Nonviscous Blade-to-Blade Flow in Rotating Blade Rows

By J. J. KRAMER,[1] CLEVELAND, OHIO

A method is developed for the blade-to-blade solution of the incompressible, nonviscous flow through rotating blade rows (including the inlet region) with or without splitter vanes. Splitter vanes are partial blades which do not extend to the inlet of the blade row. Numerical solutions are obtained for four weight flows through a centrifugal impeller without splitter vanes at one operating speed. The results are presented in a series of figures showing streamlines and relative velocity contours.

Nomenclature

THE following nomenclature is used in this paper:

A, B, C, D, E, F, G, H, I, J = points in flow field, Fig. 1

A_0, A_1, A_2, A_3, A_4 = coefficients in Equation [6]

b = stream-sheet thickness in z-direction, ft

b' = stream-sheet thickness in r-direction, ft

M = weight flow through single passage, lb/sec

r = radial distance, ft

W = ratio of relative velocity to tip speed

w = relative velocity, ft/sec

z = axial distance, ft

θ = angular co-ordinate in relative system, radians

λ = slope of trace of stream surface in axial-radial plane

ρ = fluid density, lb/cu ft

Ψ = stream function, Equation [1]

$\psi_0, \psi_1, \psi_2, \psi_3, \psi_4$ = basic solutions

ω = angular velocity of impeller, radian/sec

Subscripts

A, BC, D, E, FG, H = conditions at those stations, respectively

r = component in radial direction

tb = trailing edge of blade

tv = trailing edge of splitter vane

z = component in axial direction

θ = component in tangential direction

0 = value in basic solution of Equation [2]

1, 2 = conditions along $r = r_1$ and $r = r_2$, respectively

Introduction

The design of efficient compressors and pumps requires control of the velocity distribution of all wetted surfaces of the machine in order to prevent boundary-layer growth and separation and, in incompressible fluids, cavitation. This paper discusses a method for analyzing the flow on blade-to-blade surfaces of revolution and thus is concerned with velocity control on the blade leading edge and driving and trailing faces. An analysis method indicating these velocity distributions is necessary to warn the designer of flow conditions conducive to poor performance.

Blade-to-blade solutions of the potential flow through centrifugal compressors have been obtained by means of relaxation methods in (1 to 3).[2] In addition, a three-dimensional potential-flow solution was obtained by similar means in (4). However, all these are for impellers with inducer sections extended infinitely far upstream, or to the axis of the impeller, and thus yield no information concerning the flow behavior ahead of and at the entrance to blades of finite thickness, or on blades which are not aligned with the inlet stream.

In some applications it is desirable to insert between adjacent blades one or more splitter vanes; that is, partial blades which do not extend to the inlet of the machine. A sketch of such a blade row for a radial-inlet centrifugal impeller is shown in Fig. 1, with the splitter vane marked IJ. Splitter vanes are used to reduce the blade blockage at the inlet and at the same time to maintain reasonably low loadings on the rearward portion of the blade. This type of vane is particularly helpful in applications where blade-inlet angles are large, corresponding to low ratios of through-flow velocity to tangential velocity. No method was available for analyzing the blade-to-blade flow in blade passages with splitter vanes.

Because of the lack of a blade-to-blade analysis method which included the leading-edge region and/or permitted the presence of splitter vanes, a method covering these possibilities was de-

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

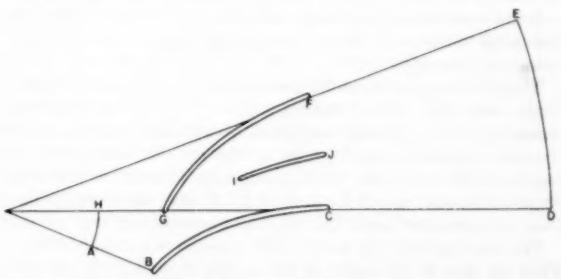


Fig. 1 Flow field for case of one splitter-vane

[1] Head, Section B, Fluid Systems Branch, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

Contributed by the Hydraulic Division and presented at a joint session with the Gas Turbine Power Division at the Semi-Annual Meeting, Cleveland, Ohio, June 17–21, 1956, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters May 29, 1956. Paper No. 56—SA-66.

veloped at the Lewis Flight Propulsion Laboratory. The method for blade rows without splitter vanes, but including the leading-edge region, and the results of the numerical application of this method to the flow in a 48-in-diam centrifugal compressor [similar to that discussed in references (5 and 6)] are reported in (7). Several solutions were obtained for various weight flows at one rotational speed. In addition, the solution in the leading-edge region was refined and reported in (8). The purposes of these numerical examples are to demonstrate the method and to show the effects on the flow in the leading-edge region caused by changes in the weight flow.

This paper contains the work reported in (7 and 8) and also extends the method to cover the case of blade rows with splitter vanes.

It is written so that those interested only in the numerical results will not find it necessary to read the Analysis section.

Analysis

The formulation of the problem and the proposed method of solution are discussed in this section.

Statement of Problem

The basic assumptions which are made concerning the physical nature of the flow determine the partial differential equation governing the flow. The assigning of proper boundary conditions to the problem then determines the particular solution of the partial differential equation.

Assumptions. The flow is assumed to be steady, incompressible, and nonviscous. The assumption of steady, nonviscous flow is customary in turbomachine-flow analyses. Several solutions have been obtained taking compressibility into account; e.g. (1 and 2). The solutions of compressible flows on blade-to-blade surfaces of revolution with subsonic relative resultant velocities can be effected by an iteration procedure of successive approximations. This procedure has been carried out in (1 and 2) and a similar procedure could be used with the method discussed herein. However, the principal contribution of this paper is the treatment of the flow in the leading-edge region and around splitter vanes. Consequently, the procedure for a compressible flow is not discussed. The fluid is assumed incompressible throughout the analysis section and in the numerical example.

The further assumption is made that the flow is constrained to a blade-to-blade surface of revolution which is symmetrical about the impeller axis. Although the flow is constrained to this surface, a variation in the thickness of the stream sheet provides a closer approximation to the actual case. The shape of the stream surface in the axial-radial plane as well as the thickness variation is defined as a function of radial position which is specified at the beginning of the solution.

These variables can be obtained from a meridional-plane solution such as that in (9 and 10). If such a solution is not available, an assumption must be made such as assuming the shape of the stream surface to be the same as the mean-blade-height line and the stream-sheet thickness, the same as the blade height in the axial direction. These assumptions were made for the numerical examples.

The rear stagnation points are assumed to be located at the blade and splitter-vane trailing edges. The Kutta condition states that, for a noncuspidate blade with a sharp trailing edge, the rear stagnation point occurs at the tip. However, for an impeller with a rounded trailing edge the location of the rear stagnation point cannot be predicted. It was necessary, therefore, to assume the location of the rear stagnation points.

The flow through the entire blade passage was obtained first. Then the flow in the region of the leading edge of the blade was determined in more detail than that obtained in the solution for the entire blade passage; the partial differential equation in the leading-edge region was solved by relaxation methods using a grid of much finer mesh. The values of the stream function obtained in the entire blade-passage solution along the boundaries of the region in which the refined solution was obtained were assumed to remain unchanged during the numerical process of obtaining the refined solution.

Differential Equation. In this analysis, the cylindrical coordinates r, θ, and z are used. Figs. 2 and 3 show the impeller for which the numerical solutions were obtained. All symbols are defined in the nomenclature. The trace of the stream surface in the axial-radial plane is given by specifying z as a function of r. The slope dr/dz of this curve is denoted by λ, Fig. 3. Thus the resultant velocity w is given by

$$w^2 = w_\theta^2 + w_r^2\left(1 + \frac{1}{\lambda^2}\right)$$

The stream function Ψ is defined by the following differential equations

$$\frac{\partial \Psi}{\partial r} = -b\rho w_\theta \quad [1a]$$

$$\frac{\partial \Psi}{\partial \theta} = rb\rho w_r \quad [1b]$$

In this paper all derivatives with respect to r shall be understood to mean derivatives with respect to r on the stream surface; that is, $\partial/\partial r$ in this paper will correspond to the bold-faced $\partial/\partial r$ of (11), in which the differential equation for the type flow considered herein is derived. With these definitions and assumptions, the differential equation of the flow becomes[3]

$$\frac{\partial^2 \Psi}{\partial r^2} + \left(\frac{1}{r} - \frac{\partial \ln b}{\partial r}\right)\frac{\partial \Psi}{\partial r} + \frac{1}{r^2}\left(1 + \frac{1}{\lambda^2}\right)\frac{\partial^2 \Psi}{\partial \theta^2} = 2\omega b\rho \quad [2]$$

This equation, together with the boundary conditions, mathematically determines the problem.

In cases where λ is zero or nearly so, it is necessary to work with a stream function defined as

$$\frac{\partial \Psi}{\partial \theta} = rb'\rho w_z \quad [3a]$$

$$\frac{\partial \Psi}{\partial z} = -b'\rho w_\theta \quad [3b]$$

which results in the following differential equation

$$\frac{(1+\lambda^2)}{r^2}\frac{\partial^2 \Psi}{\partial \theta^2} + \frac{\partial^2 \Psi}{\partial z^2} + \left(\frac{\lambda}{r} - \frac{\partial \ln b'}{\partial z}\right)\frac{\partial \Psi}{\partial z} = 2\lambda\omega b'\rho \quad [4]$$

The method will be discussed in terms of Equation [2] but a similar discussion would apply to Equation [4].

Boundary Conditions. This analysis of the flow leads to a boundary-value problem of the first kind, or a Dirichlet problem. Certain boundaries of the flow and the values of the stream function on these boundaries are specified. Furthermore, the flow is assumed to vary periodically in the circumferential direction, completing a cycle in one pitch angle, the angular distance between two adjacent blade mean lines. The rotational speed of the impeller and the weight flow through the machine also are specified. The domain of the solution is extended sufficiently far upstream and downstream so that the flow can be assumed uni-

[3] Reference (11), p. 35.

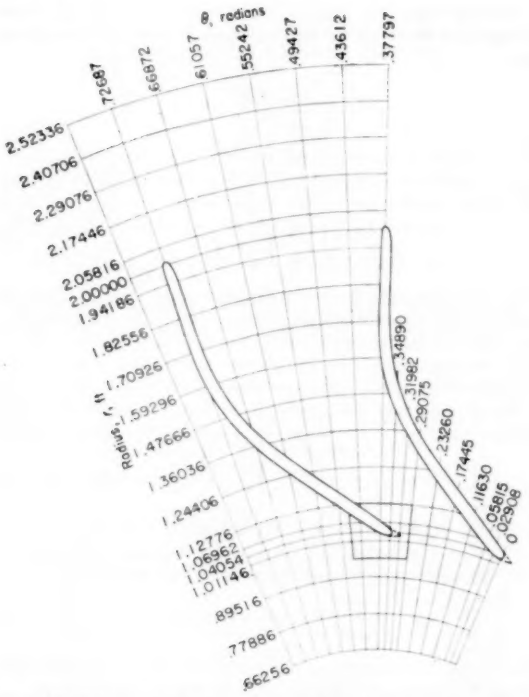


Fig. 2 Radial-tangential plane view with grid system for over-all solution

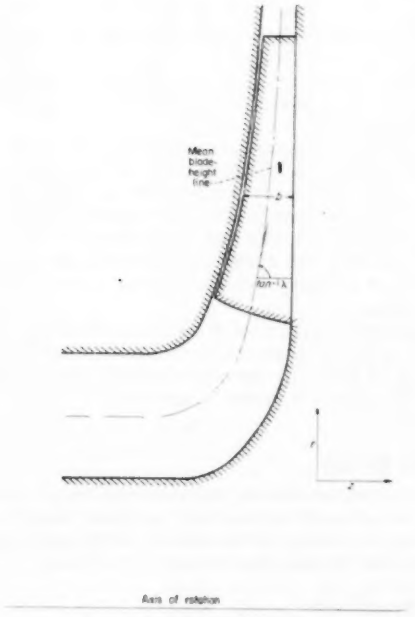


Fig. 3 Axial-radial plane view

form at the upstream and downstream boundaries. With the addition of these conditions, the problem is determined mathematically.

Method of Solution

Superposition of Basic Solutions. The differential equation (Equation [2]) is solved by a superposition of several basic solutions. These basic solutions form a set of linearly independent solutions such that all possible flows (including all rotational speeds) are expressible as linear combinations of these basic solutions. The first of these, designated ψ_0, is a solution of Equation [2] with the condition that no flow crosses the upstream and downstream boundaries and $\omega = \omega_0 \neq 0$. This solution is not necessary in the case where λ is everywhere equal to zero because in this case the differential equation is homogeneous, Equation [4].

The other basic solutions, called through-flow basic solutions, are solutions of the linear homogeneous equation obtained by equating the left side of Equation [2] to zero. Thus if

$$L = \frac{\partial^2}{\partial r^2} + \left(\frac{1}{r} - \frac{\partial \ln b}{\partial r}\right)\frac{\partial}{\partial r} + \frac{1}{r^2}\left(1 + \frac{1}{\lambda^2}\right)\frac{\partial^2}{\partial \theta^2}$$

then the through-flow basic solutions are solutions of

$$L(\Psi) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots [5]$$

Because L is a linear operator, ψ_0 plus linear combinations of the through-flow basic solutions will satisfy Equation [2] for $\omega = \omega_0$. The number of through-flow basic solutions necessary for a given problem is three greater than the number of splitter vanes between adjacent blades.

Boundary Conditions for Four Basic Solutions. For convenience, a hypothetical problem involving one splitter vane between each set of adjacent blades will be discussed. The procedure for none or more than one splitter vane will be indicated. The flow region for this hypothetical problem is represented by ABCDEFGHIJ in Fig. 1. The upstream and downstream boundaries, AH and DE, respectively, are placed sufficiently far from the blades so that flow conditions can be assumed uniform at these stations. The angular distance from A to H and from D to E is one pitch angle. For all the basic solutions, the condition that the flow is periodic about the axis of rotation with a period of one pitch angle makes it possible to obtain the solutions without a knowledge of the stream function along AB and GH. The finite-difference equation for points along these lines is obtained in the same manner as in (12). For the solution ψ_0 in which the flow is that induced only by the rotation of the impeller without any through-flow, the value of ψ along AH and DE is constant, indicating no flow crossing the upstream and downstream boundaries. The values along the blade surfaces BC, GF, and IJ are also specified as zero. The solution to Equation [2] for these boundary conditions is designated ψ_0.

The through-flow solution, that is, the flow through the stationary blade row, can be obtained from linear combinations of the four through-flow basic solutions designated ψ_1, ψ_2, ψ_3, and ψ_4. All possible flows through the stationary blade row can be represented by linear combinations of these basic solutions. It can be seen from the boundary conditions shown in Table 1 that these basic solutions are linearly independent. That these four independent through-flow-basic solutions are sufficient for the construction of all possible through flows can be seen from the following consideration.

Table 1 Boundary values for single-splitter-vane problem

Basic solution	Boundary values: At A	At D	At E	At H	Along BC	Along FG	Along IJ
ψ_0	0	0	0	0	0	0	0
ψ_1	0	0	1	1	0	1	0
ψ_2	0	−1	0	1	0	1	0
ψ_3	−1	0	1	0	0	1	0
ψ_4	0	0	1	1	0	1	1

Flow in the stationary blade row is determined when the value of ψ is determined on all bounding surfaces of the flow field. The condition of periodicity fixes the solution along AB, CD, EF, and GH. Therefore conditions must be fixed only along AH, DE, BC, IJ, and FG. The velocity is assumed to be constant at stations AH and DE. Hence the stream function varies linearly from A to H and from D to E. Thus conditions along AH and DE are determined by ψ-values at A, H, D, and E. Because A and D are spaced one pitch angle from H and E, respectively, $\psi_H - \psi_A$ is equal to $\psi_E - \psi_D$. Thus conditions along AB and DE are determined by specifying ψ_A and ψ_D for a specified difference $\psi_H - \psi_A$. The ψ-values along BC and FG are constants differing by an amount equal to $\psi_H - \psi_A$. Thus only one of the values ψ_{BC} and ψ_{FG} is independent. The value of ψ is the same along both driving and trailing faces of IJ. Thus ψ-values must be determined at seven stations (A, H, D, E, BC, FG, and IJ). However, only four can be chosen which are independent. The four values chosen in this problem are A, D, FG, and IJ. It can be seen from Table 1 that any ψ-value can be obtained at these stations in a solution formed by linear combinations of ψ_1, ψ_2, ψ_3, and ψ_4. Because a solution of Equation [5] remains a solution when changed by a multiplicative or additive constant, the specification of the difference $\psi_H - \psi_A$ mentioned in the discussion is no restriction.

For the case of no splitter-vanes solution ψ_4 could be eliminated so that the final solution for flow in the rotating-blade row would be effected by superposition of four basic solutions. The numerical example presented in this paper is an instance of such a procedure.

For cases of more than one splitter vane the boundary conditions would be chosen in a manner analogous to that for the one-splitter-vane case. These boundary conditions must be linearly independent and form a basis for constructing the desired solution.

Coefficients of ψ_0, ψ_1, ψ_2, ψ_3, and ψ_4 in Linear Combinations. The final solution Ψ for any weight flow or rotational speed will be obtained from an equation of the form

$$\Psi = A_0\psi_0 + A_1\psi_1 + A_2\psi_2 + A_3\psi_3 + A_4\psi_4 \quad [6]$$

The coefficients A_0, A_1, A_2, A_3, and A_4 are determined by the specification of five independent physical conditions: (*a*) The rotational speed; (*b*) the weight flow; (*c*) the location of the rear stagnation point on the blade; (*d*) the location of the rear stagnation point on the splitter vane; and (*e*) irrotationality of the absolute flow.

The coefficient A_0 is determined by the rotational speed and is given by

$$A_0 = \frac{\omega}{\omega_0} \quad [7]$$

that is, A_0 is the ratio of the rotational speed ω for the desired solution to that used in obtaining the basic solution ω_0.

The change in Ψ across one blade passage is equal to the weight flow through a single passage M. Therefore

$$A_1 + A_2 + A_3 + A_4 = M \quad [8]$$

The rear stagnation points of the blade and the splitter vane are assumed to be at the trailing edges. Thus at the trailing edges

$$w_\theta = 0$$

or

$$\left(\frac{\partial\Psi}{\partial r}\right)_{tb} = 0 \quad [9]$$

and

$$\left(\frac{\partial\Psi}{\partial r}\right)_{tv} = 0 \quad [10]$$

where the subscripts tb and tv denote trailing edge of blade and trailing edge of splitter vane, respectively. These derivatives are expressed in finite-difference form for the grid points at the blade trailing edge and at the splitter-vane trailing edge and with Equation [6] yield two linear relations in A_0, A_1, A_2, A_3, and A_4.

The absolute flow is irrotational, so that, if r_1 and r_2 are radial stations upstream of the blade row, the following equation holds

$$\int_0^{2\pi} (w_{\theta,2} + \omega r_2) r_2 d\theta - \int_0^{2\pi} (w_{\theta,1} + \omega r_1) r_1 d\theta = 0 \quad [11]$$

where the subscripts 1 and 2 indicate values along the lines $r = r_1$ and $r = r_2$, respectively. If r_1 is chosen equal to the value of r at the upstream boundary, Equation [11] becomes

$$\int_0^{2\pi} r_2 w_{\theta,2} d\theta = -2\omega\pi r_2^2 \quad [12]$$

because $w_{\theta,1}$ is equal to $-\omega r_1$. When the stream-function definition, Equation [1*a*], is introduced, Equation [12] becomes

$$\int_0^{2\pi} \frac{\partial\Psi}{\partial r}\, d\theta = 2\pi\omega\rho b_2 r_2 \quad [13]$$

Equation [13] can be integrated numerically to yield a linear relation in A_0, A_1, A_2, A_3, and A_4.

Equations [7], [8], [9], [10], and [13] form a system of five simultaneous independent linear equations in five unknowns A_0, A_1, A_2, A_3, and A_4.

From the previous discussion it can be seen that as the number of splitter vanes increases or decreases so also the number of coefficients and the number of linear relations among the coefficients increase or decrease correspondingly in a one-for-one manner.

Numerical Method of Obtaining Basic Solutions. The region of solution is covered with a network of grid lines the intersections of which form grid points, as shown in Fig. 2, for the numerical example computed.

The solution for a given set of boundary conditions of the differential equation is obtained at each of these grid points by solving the set of linear simultaneous equations obtained when the differential equation is written in finite-difference form for each grid point.

Numerical Example

The previously outlined method was applied in order to analyze the flow in a 48-in-diam radial-inlet centrifugal impeller. A description of the geometry of the impeller and the operating conditions for which the analysis was carried out, follows.

Application of Method

Geometry of Impeller. The impeller investigated was a 48-in-tip-diam radial-inlet centrifugal impeller having 18 blades, similar to that discussed in (5 and 6). The sharp leading edge and blunt trailing edge were rounded as shown in Fig. 2 because of practical computing considerations. The blade co-ordinates are given in

Table 2. The solution was obtained on the surface generated by rotating the mean blade-height line about the axis of rotation. This line was approximated by the following function

$$z = \frac{-0.041456}{r - 0.40828} + \text{const} \quad \ldots\ldots\ldots\ldots [14]$$

The streamline spacing in the axial-radial plane is not known. Therefore the stream-sheet thickness b in the z-direction was approximated by the blade height in the z-direction. This parameter was approximated by the following function

$$b = 0.07208 + 1.01517\, e^{-1.84601r}$$

The parameter λ is equal to dr/dz of the stream-surface trace in the axial-radial plane and from Equation [14] is given by

$$\frac{1}{\lambda} = \frac{0.041456}{(r - 0.40828)^2}$$

Table 2 Modified blade co-ordinates

Driving face		Trailing face	
r, ft	θ, radians	r, ft	θ, radians
1.0405	0.34256	1.0285	0.34890
1.0521	0.34890	1.0352	0.37797
1.0696	0.37224	1.0405	0.38604
1.0740	0.37797	1.0518	0.40705
1.1190	0.43612	1.0696	0.43273
1.1278	0.44633	1.0971	0.46520
1.1711	0.49427	1.1278	0.49939
1.2324	0.55242	1.1511	0.52335
1.2441	0.56242	1.2144	0.58150
1.3095	0.61057	1.2441	0.60413
1.3604	0.63995	1.2974	0.63965
1.4262	0.66872	1.3525	0.66872
1.4767	0.68476	1.3604	0.67227
1.5930	0.70726	1.4283	0.69780
1.7093	0.71497	1.4767	0.71113
1.8256	0.71601	1.5600	0.72687
1.9419	0.71601	1.5930	0.73041
2.0000	0.72080	1.7093	0.73583
2.0123	0.72687	1.8256	0.73416
		1.9419	0.73312
		2.0000	0.73145
		2.0123	0.72687

Operating Conditions. Four solutions were obtained corresponding to four weight flows at a tip speed of 700 ft per sec. These four weight flows, which correspond to those of (5), are as follows:

Case A.....14.00 lb per sec
Case B.....26.25 lb per sec
Case C.....32.10 lb per sec
Case D.....44.00 lb per sec

Numerical Procedure. The region of solution was covered with the grid as shown in Fig. 2. A five-point system was used to express the derivatives in the finite difference equation at each grid point corresponding to the differential Equation [2]. The solution of the set of n simultaneous linear equations (where n is the number of grid points) was obtained on high-speed digital computers by the matrix method outlined in (12). Since there were four basic solutions, four sets of n simultaneous linear equations were solved by this process. The solutions obtained for the entire flow field were called the over-all solutions.

In addition, the solutions in the leading-edge region were refined by solving the differential equation of flow by relaxation methods on a grid of much finer mesh size shown in Fig. 4. For this refinement the boundary values were obtained from the over-all solutions. The residuals of the relaxation process were reduced to values indicating zero change in the fifth decimal place of the stream function.

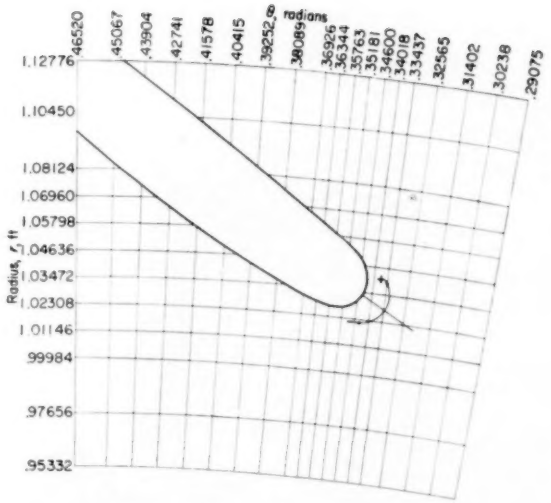


Fig. 4 Grid system for refined solution

Results and Discussion of Numerical Example

The results of the solutions obtained by the application of the previously outlined methods are presented in Figs. 5 to 12, which show streamlines and constant relative-velocity contours. The over-all solution for the entire blade passage as obtained by the matrix method is shown in the (*a*) part of each figure and the refined solution for the leading-edge region in the (*b*) part. Figs. 5 to 12 are projections on the $r\theta$-plane; that is, the curvature of the stream surface in the axial-radial plane is neglected.

Streamlines

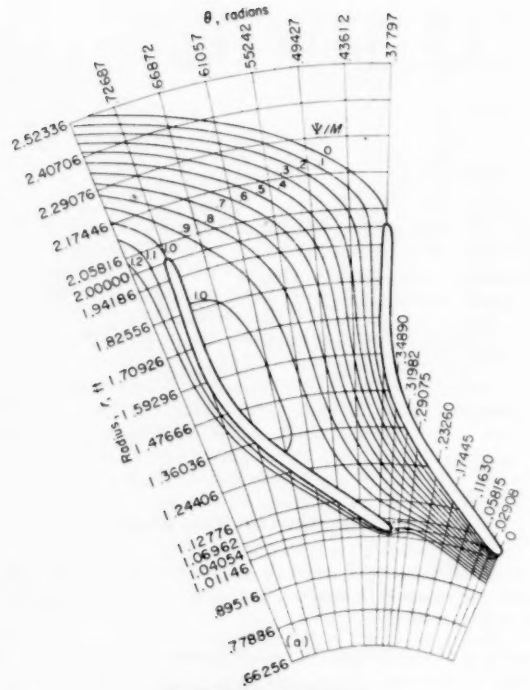
The distribution of stream function is shown by means of contours of constant stream-function ratio Ψ/M in Figs. 5 to 8 for the four weight flows investigated. The impeller tip speed was 700 ft per sec for all four cases.

Case A, Fig. 5, corresponds to the incipient surge weight flow for the experimental case (5). A large eddy attached to the driving face of the blade extends from $r \sim 1.31$ to $r \sim 1.84$ ft and almost one third of the distance across the passage between blades at its widest point. The major part of the flow is concentrated in the region near the trailing face, while the eddy and other relatively low-momentum fluid occupy half the channel. The inlet stagnation point occurs on the driving face of the blade at $r \sim 1.05$ ft.

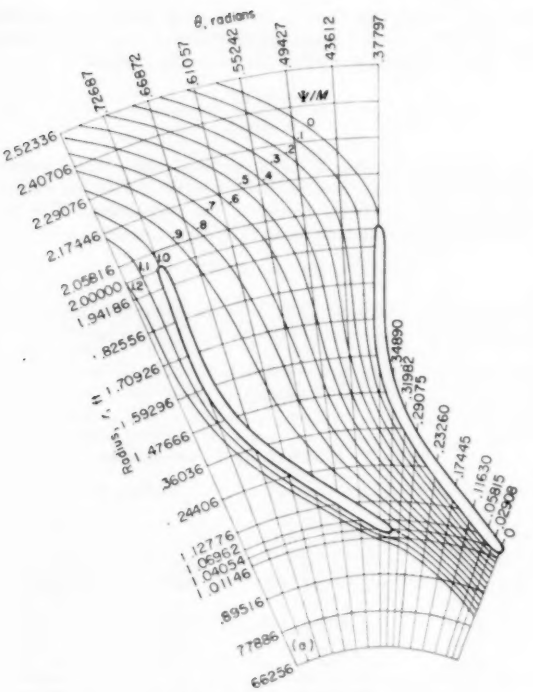
The weight flow for case B, Fig. 6, is sufficiently high to eliminate the eddy on the driving face of the blade. However, a fairly large concentration of low-momentum air is still present, so that halfway through the impeller one half of the fluid occupies more than two thirds the available flow area.

The streamline pattern for case C, Fig. 7, is similar to that for case B because of the small change in weight flow.

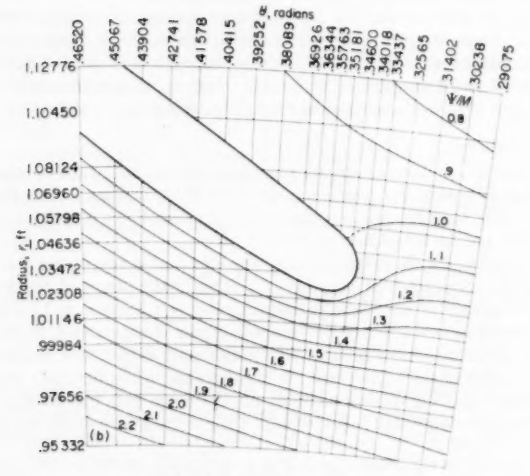
In the investigation reported in (5), the weight flow corresponding to case D, Fig. 8, represented the maximum weight flow attainable experimentally at a tip speed of 700 ft per sec. The flow is distributed across the passage more nearly uniformly than in the other examples. The flow ceases to be perfectly guided at $r \sim 1.5$ ft, as occurred for all other weight flows. The slip factor, the ratio of the mass-averaged absolute tangential



(*a*) Entire blade passage

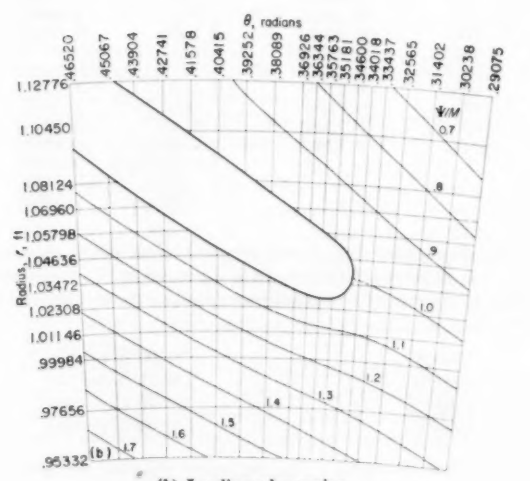


(*a*) Entire blade passage



(*b*) Leading-edge region

Fig. 5 Streamlines for Case A



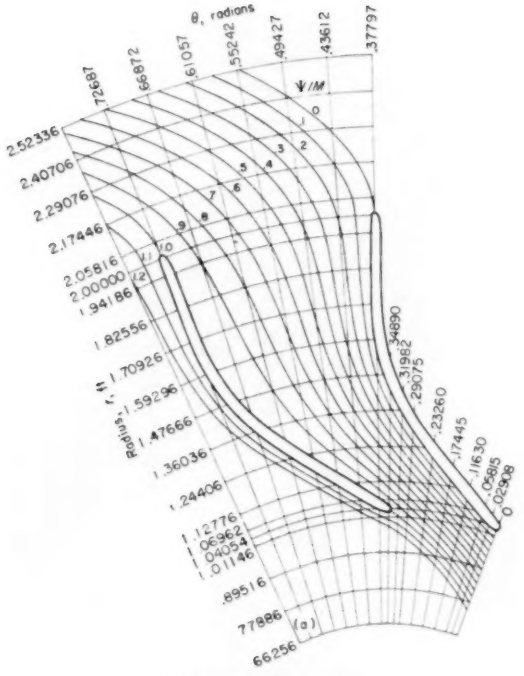
(*b*) Leading-edge region

Fig. 6 Streamlines for Case B

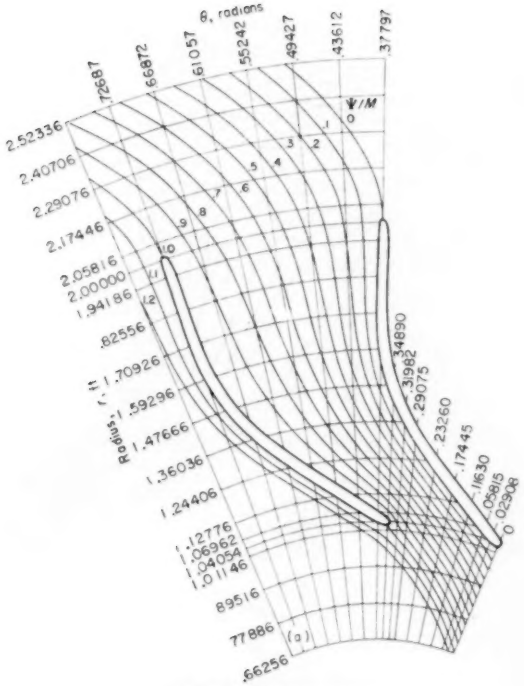
velocity at the tip to the absolute tip speed, decreased with increasing weight flow from 0.874 for case A to 0.859 for case D. The inlet stagnation point occurs on the trailing face at $r \sim 1.028$ ft. Thus the stagnation point shifts from the driving to the trailing face as the weight flow increases from 14 to 44 lb per sec.

Relative Velocity

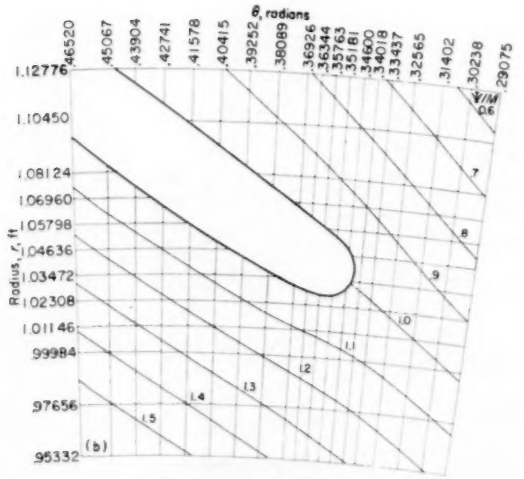
Contours of constant relative velocity ratio W (relative velocity divided by tip speed) are plotted in Figs. 9 to 12 for the four weight flows investigated. In the figures showing the entire blade passage—(*a*) parts—the velocities near the leading edge are not shown. Reference should be made to the figures showing the leading-edge region only—(*b*) parts. For case A, Fig. 9, at $r \sim 1.31$ ft on the driving face, a stagnation point occurs where the eddy begins to form. Velocities are low along the entire driving face with negative velocities in the blade-surface eddy region. A rapid acceleration followed by a less rapid deceleration occurs on the leading edge and the trailing face because of the positive angle of attack (inlet flow directed toward the driv-



(*a*) Entire blade passage

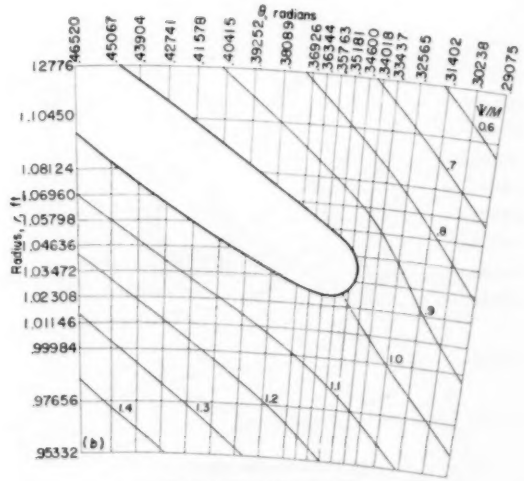


(*a*) Entire blade passage



(*b*) Leading-edge region

Fig. 7 Streamlines for Case C



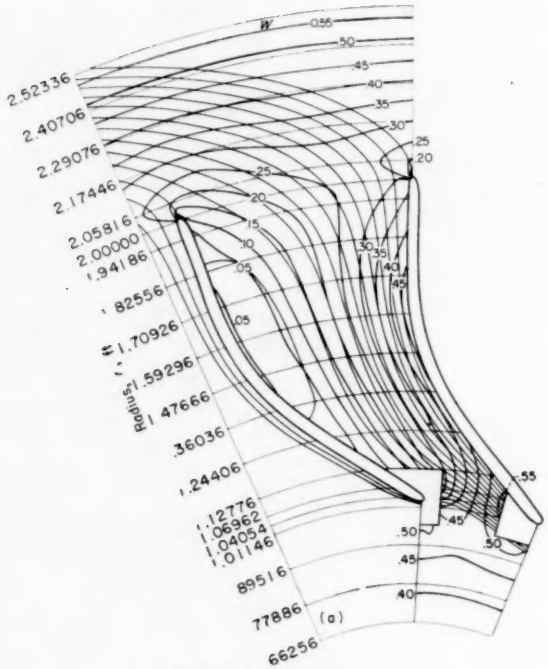
(*b*) Leading-edge region

Fig. 8 Streamlines for Case D

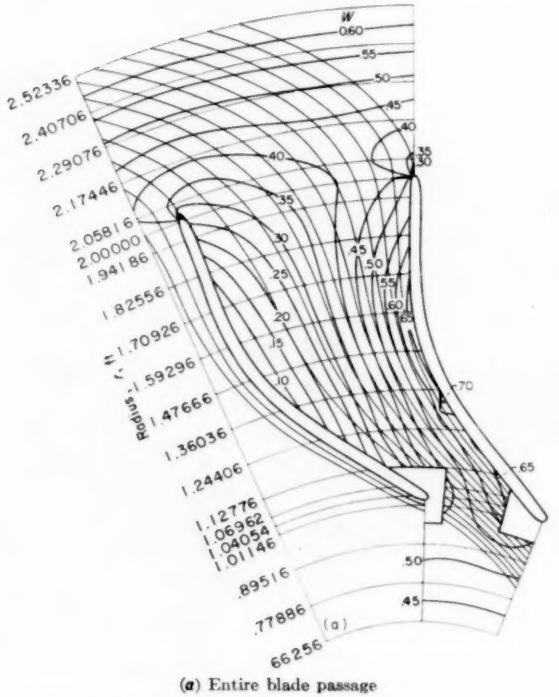
ing face). This acceleration and deceleration shift around to the driving face as the weight flow increases.

For case B, Fig. 10, downstream of the leading-edge region the velocity along the trailing face is nearly constant (except for a small acceleration and deceleration at $r \sim 1.3$ ft) to $r \sim 1.7$ ft. In the leading-edge region small local decelerations occur on both the driving and trailing faces with the one on the trailing face being the larger. Flow conditions seem to be the best for this case corresponding to 26.25 lb of air per sec. At $r \sim 1.3$ ft on the trailing face an acceleration occurs followed by a rapid deceleration in cases B, C, and D. This velocity peak is caused by the beginning of more rapid blade curvature at that point and becomes more pronounced as the weight flow increases.

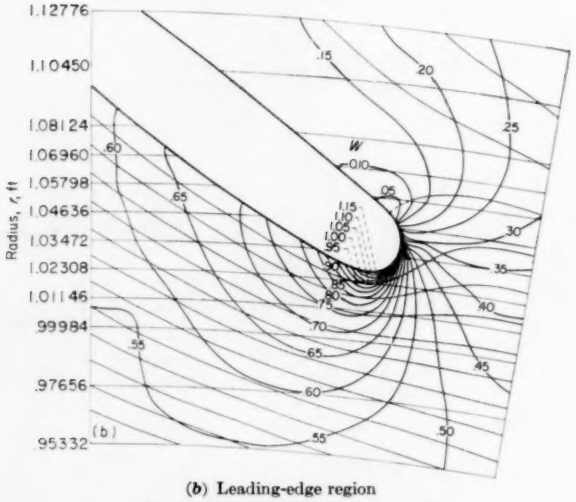
For case C, Fig. 11, a larger deceleration occurs on the driving face than for case B. Decelerations are probably more serious on the driving face than on the trailing face because the low-momentum air caused by the deceleration aggravates the secondary-flow conditions. These secondary flows transport the low-momentum fluid on the driving face to the trailing face. This type of motion is discussed in more detail in (13).



(*a*) Entire blade passage

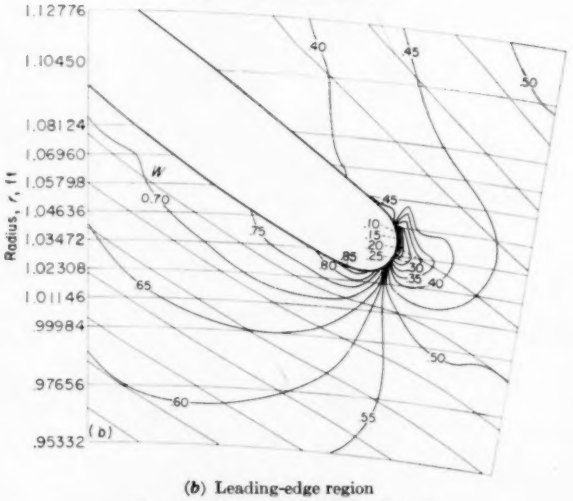


(*a*) Entire blade passage



(*b*) Leading-edge region

Fig. 9 Velocity Contours for Case A



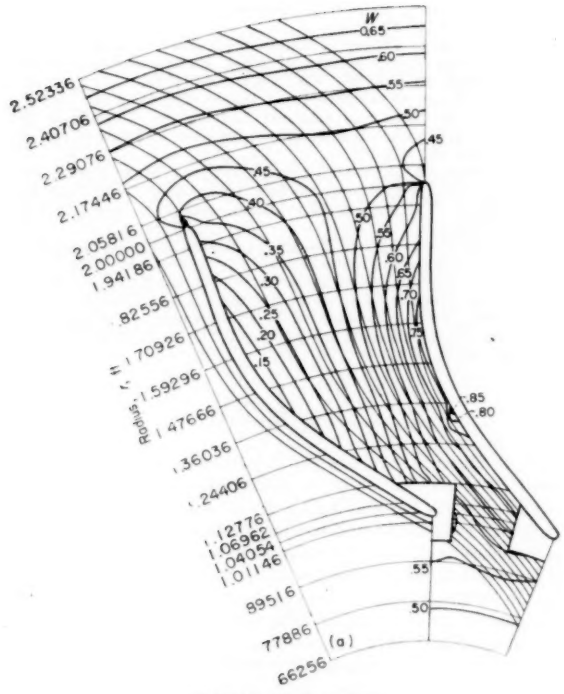
(*b*) Leading-edge region

Fig. 10 Velocity Contours for Case B

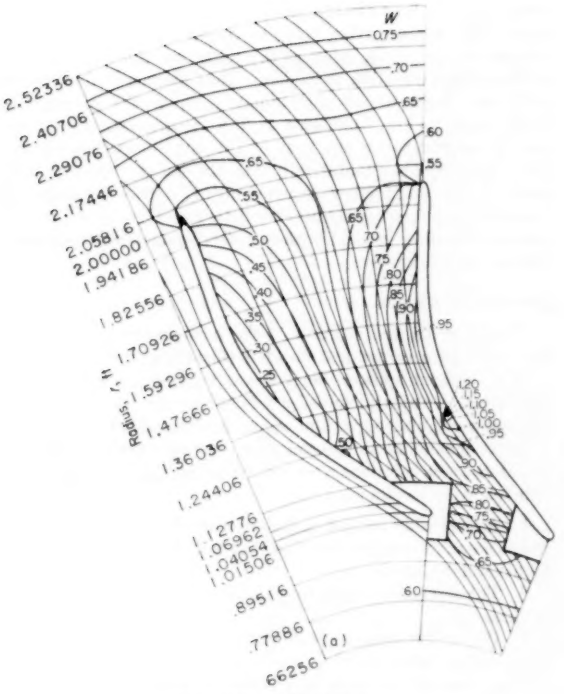
For case D, Fig. 12, a large deceleration occurs on the driving face just downstream of the leading edge. This deceleration is about the same size as that which occurred on the trailing face in case A. The deceleration is probably more serious on the driving face because of its contribution to the build-up of secondary flows. Also, at sufficiently high weight flows the separation following a rapid deceleration will induce choking before the theoretical maximum weight flow is attained.

Mean Angle of Attack

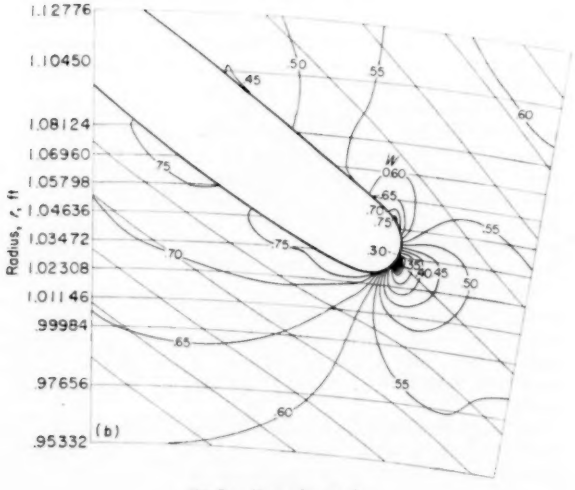
The approximate mean angle of attack, that, is the angle between the mass-averaged flow direction at the inlet and the tangent to the blade mean line, was computed from the rotational speed and the average inlet velocity. The average inlet velocity was computed from the weight flow and the annular area. Two values were used for the annular area: (*a*) The total annular area with no blade blockage assumed; and (*b*) the total annular area minus the blockage caused by the blades. The thickness of the blades used in the latter computation was that at the 1.04-ft radius, which was approximately the radius at which maximum blade thickness in the tangential direction occurred. The mean angle of attack across the passage at the 1.04-ft radius was also computed from the exact solution. These are compared in Table 3.



(*a*) Entire blade passage

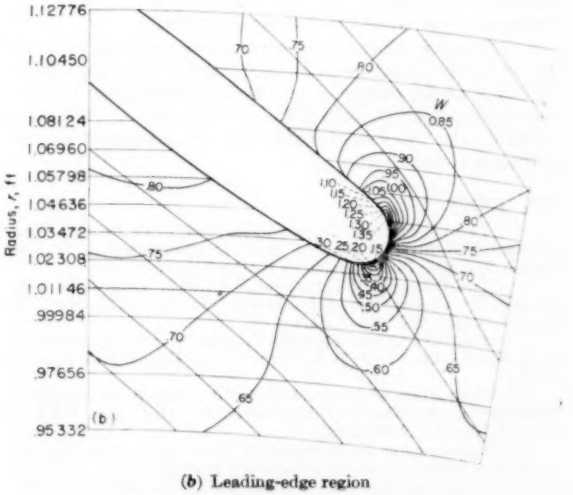


(*a*) Entire blade passage



(*b*) Leading-edge region

Fig. 11 Velocity Contours for Case C



(*b*) Leading-edge region

Fig. 12 Velocity Contours for Case D

Table 3 Variation of various mean angles of attack with weight flow

Weight flow, lb/sec	Mean angle of attack, based on— Unblocked annulus, deg	Blocked annulus, deg	Exact solution, deg
14	12.3	9.2	0
26.25	0.1	−4.6	−5.3
32.10	−4.9	−10.0	−9.0
44	−13.6	−19.0	−15.4

The sign convention for the angle of attack is such that a positive angle of attack indicates flow directed toward the driving face of the blade. From the comparison of these angles of attack, it is apparent that the mean angle of attack is best predicted by basing the calculations on the annular area with blade blockage considered. The poor agreement between the mean angle of attack of the exact solution and that based on the blocked-inlet annular area at the lowest weight flow is probably caused by the eddy. It appears that inlet flow aligned with the driving face

results in good flow conditions in the leading-edge region. The blade angle of the driving face just downstream of the rounded leading edge is 57 deg, whereas the angle between the mean line and the radial direction is 62 deg. Thus for case B the average inlet flow angle would approximately equal the driving-face blade angle. Flow conditions in the leading-edge region for case B seemed to be the best of the conditions investigated. This result agrees qualitatively with the design procedure suggested in (13) and further discussed in (14). These leading-edge contours are characterized by very little curvature of the driving face so that flow aligned with the driving face would produce little or no deceleration.

Summary of Results

A method for the solution of the incompressible nonviscous flow through a centrifugal impeller (including the inlet region) with or without splitter vanes was developed and applied to a 48-in-diam centrifugal impeller. Solutions for the entire blade passage were obtained for four weight flows ranging from incipient surge to maximum as determined by actual impeller tests. In addition, these solutions were refined in the leading-edge region. The following results were noted:

1 A large eddy formed on the driving face of the blade at the incipient surge weight flow but was not present for the three higher weight flows.

2 The slip factor varied from 0.874 to 0.859 as the weight flow increased.

3 For weight flows of 26.25, 32.10, and 44 lb per sec, a local acceleration followed by a rapid deceleration occurred on the trailing face of the blade at a radius of about 1.3 ft; that is, where the blade began to curve more rapidly.

4 The mean angle of attack was best predicted by basing the approximate computation on the weight flow, the tip speed, and the annular area minus the blockage of the blades.

5 Minimum velocity gradients around the blade nose occurred for the weight flow corresponding to a mean angle of attack of −4.6 deg computed from blade speed and the upstream radial-axial velocity for which blade blockage has been taken into account. For this condition the inlet flow was aligned with the driving face of the blade.

Bibliography

1 "Two-Dimensional Compressible Flow in Centrifugal Compressors With Straight Blades," by J. D. Stanitz and G. O. Ellis, NACA Rep. 954, 1950 (supersedes NACA TN 1932).

2 "Two-Dimensional Compressible Flow in Centrifugal Compressors With Logarithmic-Spiral Blades," by G. O. Ellis and J. D. Stanitz, NACA TN 2255, 1951.

3 "Two-Dimensional Flow on General Surfaces of Revolution in Turbomachines," by J. D. Stanitz and G. O. Ellis, NACA TN 2654, 1952.

4 "Comparison of Two and Three-Dimensional Potential-Flow Solutions in a Rotating Impeller Passage," by G. O. Ellis and J. D. Stanitz, NACA TN 2806, 1952.

5 "Experimental Investigation of Flow in the Rotating Passages of a 48-Inch Impeller at Low Tip Speeds," by D. J. Michel, Ambrose Ginsburg, and John Mizisin, NACA RM E51D20, 1951.

6 "Theoretical Analysis of Incompressible Flow Through a Radial-Inlet Centrifugal Impeller at Various Weight Flows. I—Solution by a Matrix Method and Comparison With an Approximate Method," by V. D. Prian, J. J. Kramer, and Chung-Hua Wu, NACA TN 3448, 1955.

7 "Theoretical Analysis of Incompressible Flow Through a Radial-Inlet Centrifugal Impeller at Various Weight Flows. II—Solution in Leading-Edge Region by Relaxation Methods," by J. J. Kramer, NACA TN 3449, 1955.

8 "Two Axial-Symmetry Solutions for Incompressible Flow Through a Centrifugal Compressor With and Without Inducer Vanes," by G. O. Ellis and J. D. Stanitz, NACA TN 2464, 1951.

9 "Method of Analysis for Compressible Flow Through Mixed-Flow Centrifugal Impellers of Arbitrary Design," by J. T. Hamrick, Ambrose Ginsburg, and W. M. Osborn, NACA Rep. 1082, 1952. (supersedes NACA TN 2165).

10 "A General Theory of Three-Dimensional Flow in Subsonic and Supersonic Turbomachines of Axial, Radial, and Mixed-Flow Types," by Chung-Hua Wu, NACA TN 2604, 1952.

11 "A Theory of the Direct and Inverse Problems of Compressible Flow Past Cascade of Arbitrary Airfoils," by Chung-Hua Wu and Curtis A. Brown. *Journal of the Aeronautical Sciences*, vol. 19, March, 1952, pp. 183–196.

12 "Study of Three-Dimensional Internal Flow Distribution Based on Measurements in a 48-Inch Radial-Inlet Centrifugal Impeller," by J. T. Hamrick, John Mizisin, and D. J. Michel, NACA TN 3101, 1954.

13 "Die Stromung um die Schaufeln von Turbomachinen," by F. Weinig, Johann Ambrosius Barth, Leipzig, Germany, 1935.

14 "Effect of Blade-Thickness Taper on Axial-Velocity Distribution at the Leading Edge of an Entrance Rotor-Blade Row With Axial Inlet, and the Influence of This Distribution on Alignment of the Rotor Blade for Zero Angle of Attack," by J. D. Stanitz, NACA TN 2986, 1953.

Discussion

G. O. Ellis.[4] The author and the National Advisory Committee for Aeronautics are to be congratulated for this important contribution to the growing library of technical papers dealing with flow conditions inside rotating blade rows. Such theoretical approaches have been very valuable in the transition of centrifugal-compressor design from an art to a science. The science is still in its infancy and experimental evaluation of the theory and the determination of critical limits of flow are still needed. It is hoped that the NACA will continue its efforts along this line.

In connection with the author's discussion of the effect of the blade thickness on the mean angle of attack, it should be pointed out that the thickness and taper of the blade, in the presence of an axial-radial turn, causes a shift in the meridional distribution of the flow so that the difference in the average meridional component of velocity between the blocked and unblocked passage cannot be fully accounted for by the local blockage alone, as inferred from the author's two-dimensional solutions.

This is demonstrated by the following two examples of flow in an annular passage. The meridional velocities were determined in an annulus having no vanes and are shown in Fig. 13 herewith. For the second example, the average meridional velocities were determined in the same annular passage but with nonloaded blades, i.e., blades so aligned with the flow that the only effect was assumed to be due to the thickness of the blades. A plot of these velocities is shown in Fig. 14.

Note that near the hub, where the vane occupies about 30 per cent of the local area, the meridional velocity is not significantly changed by the addition of vanes while near the shroud, where the vane occupies only about 12 per cent of the local area, the meridional velocity is increased 18 per cent by the addition of vanes.

Two suggestions are offered concerning the basic solutions: (1) General solutions can be obtained using one less basic solution than proposed by the author if Equation [13] of the paper is used to supply the upstream boundary conditions for each of the basic solutions. Since the location of station 2 (the station at which Equation [13] is applied) can be taken any place upstream from the blade inlet, let it be taken at the upstream boundary. Since conditions are specified as uniform here, $\partial\psi/\partial r$ is a constant and so Equation [13] can be integrated and $\partial\psi/\partial r$ evaluated as follows

$$\left.\frac{\partial\psi}{\partial r}\right)_2 = \omega\rho b_2 r_2 \qquad [15]$$

[4] Research and Development Division, Carrier Corporation, Syracuse, New York.

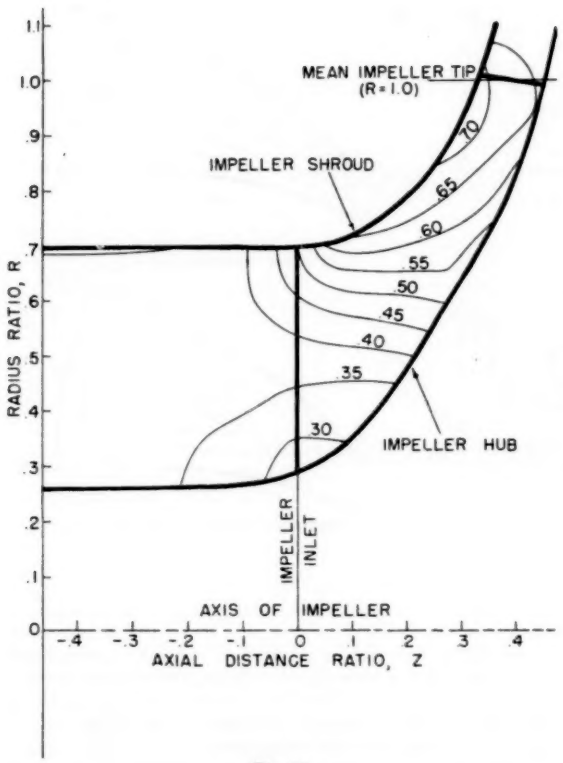


Fig. 13

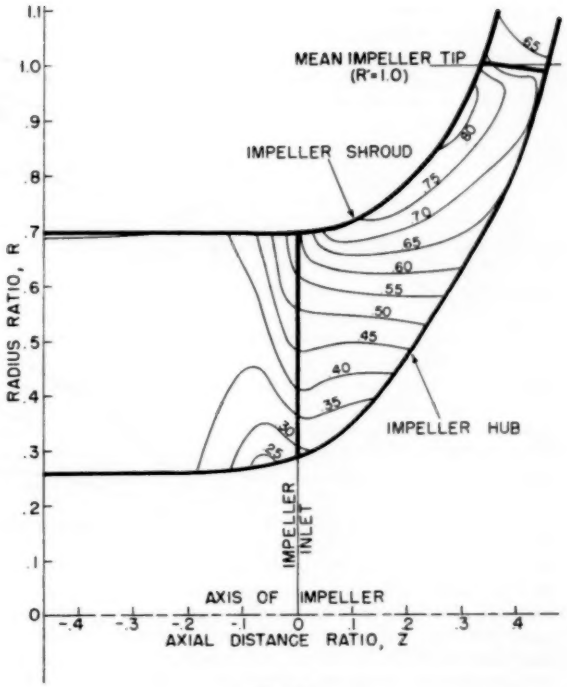


Fig. 14

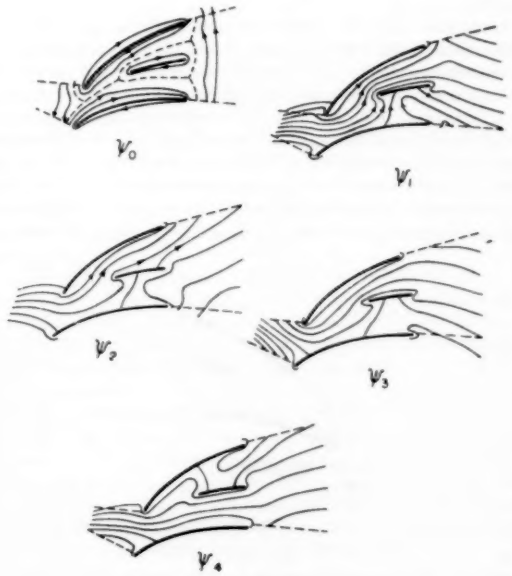


Fig. 15

Equation [15] can be used as the specified boundary condition for each of the basic solutions with $\omega = \omega_0$ for the nonhomogeneous solutions, and equal to zero for the homogeneous solutions.

If solutions are obtained on a high-speed digital computer the number of basic solutions required is perhaps unimportant, but if the solutions are obtained by manual relaxation techniques the elimination of one solution should be time-saving.

The second suggestion concerns the selection of boundary conditions to give the smoothest possible flow patterns for the basic solutions. If analytic expressions could be obtained for the basic solutions, the streamline configurations would be of little significance. Finite-difference approximations are involved in the present solutions, however, which assume that the stream-function distribution can be approximated by a polynomial expression over a given distance. As the flow becomes more complex, either a higher order polynomial must be used to approximate the stream-function distribution over the given distance or the distance must be reduced (assuming no loss in accuracy is to be allowed). Stated more simply: As the flow becomes more complex, either the finite-difference expression must become correspondingly more complex, or a larger number of grid points, i.e., smaller grid spacings, must be used to obtain a solution. If solutions are obtained by machine, this means additional storage space is needed and if they are obtained by manual relaxation more time is needed.

Streamlines, for the basic solutions proposed by the author for the case of one splitter vane, probably would look something like the sketches in Fig. 15 of this discussion. The flow becomes increasingly complex as more splitter vanes are added as seen in Fig. 16.

There may be an advantage in selecting the basic boundary conditions as shown in Fig. 17. The boundary conditions for the nonhomogeneous solution, ψ_0, are selected to represent flow at some intermediate condition of the range which is to be investigated. No difficulty should be encountered in establishing these boundaries. M_0 is determined by the specified weight flow. The downstream boundary can be established from con-

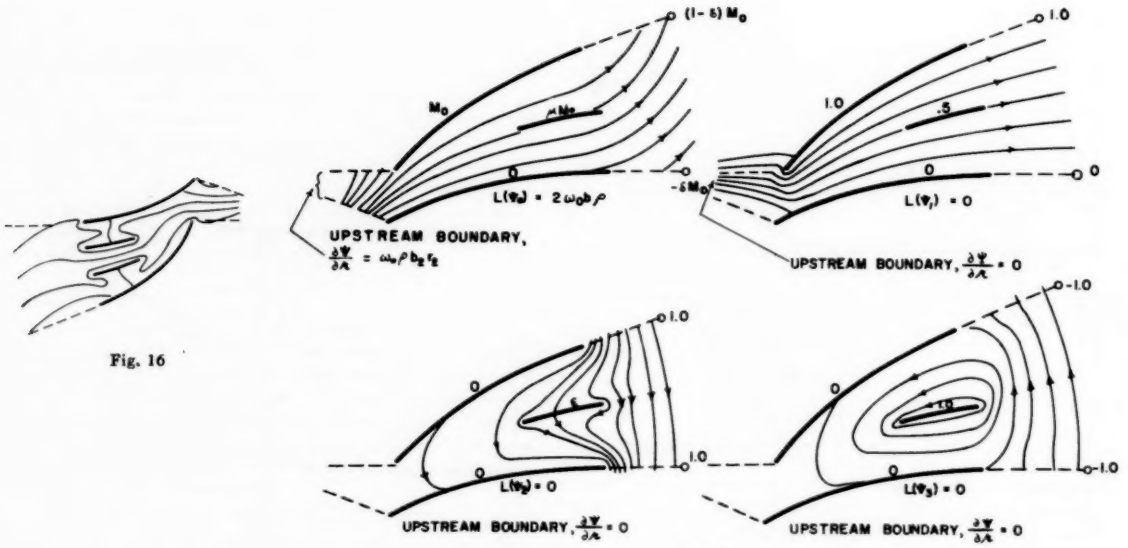


Fig. 16

Fig. 17

sideration of an estimated slip factor and the specified rotational speed of the impeller so that only small corrections are needed to satisfy the Kutta conditions. The upstream boundary condition has already been discussed. The value of ψ attached to the splitter vane can only be approximated, but experience suggests a value between 0.6 and 0.8 of M_0 would not disrupt the flow too badly.

Variations in weight flow can be accounted for by addition or subtraction of the homogeneous solution ψ_1, and corrections for the Kutta condition at the tips of the main vanes and the splitter vanes are obtained from the remaining two basic solutions. It is seen, of course, that the streamline picture for ψ_2 and ψ_3 are as complex as those of the author, but it should be pointed out that these solutions represent second-order corrections which are added to a combination of the other two solutions. Thus solutions obtained using the same grid system and finite difference approximations as used to obtain ψ_0 will be sufficiently accurate.

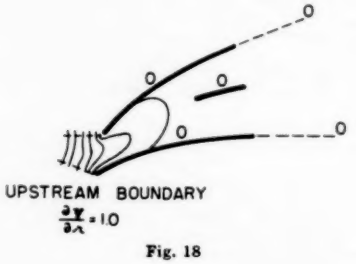


Fig. 18

A basic solution which can be used to introduce the effects of variable prewhirl is shown in Fig. 18. In this solution, all boundary values are zero except the upstream value where $\partial\psi/\partial r$ is given an arbitrary value, say 1.

J. T. Hamrick.[5] The author is to be congratulated on the presentation of an excellent and timely paper. In presenting a solution of the flow at the inlet, to a centrifugal impeller, the author fills a conspicuous gap in the literature on centrifugal machinery. To this reviewer's knowledge, a solution comparable to this one (excluding the NACA reports on which the paper was based) has not been presented. This paper points up the need for careful analysis of additional leading-edge shapes, and in particular, those of reference (14) of the paper. Solutions are especially desirable for application to centrifugal-pump impellers where acceleration to high velocities at the leading edge can result in destructive cavitation.

Another desirable objective is the attainment of a rapid method of analysis of flow at the leading edge such as those which exist for application inside rotating passages where channel flow is more nearly approximated.

There is one remaining significant gap in potential flow solutions and that is the one for impellers with splitter vanes. At present, the advantage of splitter vanes is questionable because of a lack of knowledge on how to design them. In this reviewer's experience, they have been useful in reducing the blade loading sufficiently to eliminate flow instability, but have resulted in no gain in efficiency. In fact, they usually reduce the efficiency of a well-designed impeller by as much as a couple of points. Potentially, they should raise the efficiency as well as lower manufacturing costs. The method given by the author for analysis with splitter vanes appears time-consuming. Would he care to give an estimate as to the amount of time required for such a solution?

F. S. Weinig.[6] The author should be commended for the diligence with which he treated his problem in general and for the care with which details have been worked out and represented by graphs. What may have impressed me mostly is the behavior of the flow near the inlet. The example shows again that the shape of the leading edge deserves much attention if local or permanent separation should be avoided, but even more so if danger of cavitation has to be minimized or at least controlled as far as possible.

[5] Engineering Specialist, Thompson Products, Inc., Cleveland, Ohio. Mem. ASME.

[6] Aerodynamicist, Component Development Section, AGT Development Department, General Electric Company, Cincinnati, Ohio. Mem. ASME.

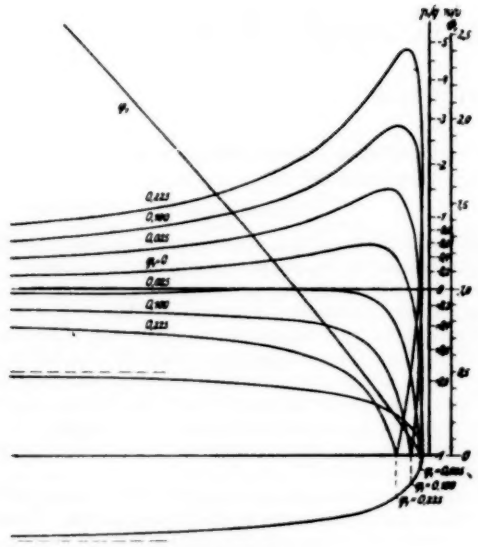


Fig. 19

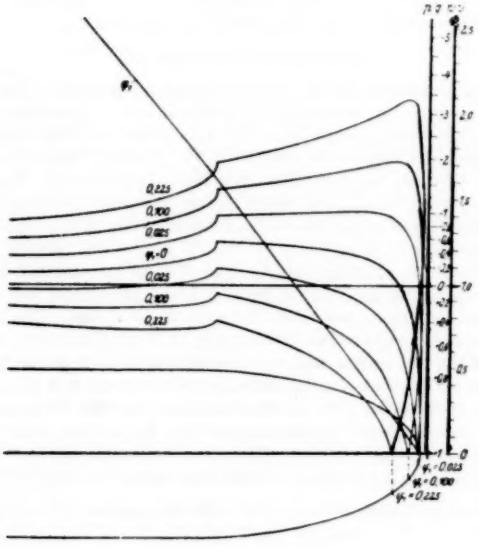


Fig. 20

What influence the shape of the leading edge has on the flow at comparable changes of operational conditions may be observed from Figs. 19 and 20 of this discussion.[7] The first shows the pressure distribution along a leading edge as produced by superimposition of a source and a parallel flow, as used at least similarly by Mr. Kramer; the second on a leading edge of an almost semi-ellipsoidal shape continued by blade of constant thickness. At comparable off-design conditions, the first yields a velocity ratio, e.g., of 2.42 against the average velocity; the second only 2.08. This difference may be quite decisive for separation and for cavitation or their avoidances.

[7] Taken from "Zur Frage der Abrundung und Zuschärfung umströmter Kanten," by F. Weinig, *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 13, 1933, p. 224.

Author's Closure

The author wishes to thank the reviewers for their efforts spent in commenting on the paper. Mr. Ellis is correct in stating and showing that the angle of attack cannot be predicted on the basis of blade blockage alone. The encouraging aspect of the results presented in the paper is that the angle of attack as found in a two-dimensional blade-to-blade solution is approximated reasonably well by a simple calculation based on blade blockage.

In regard to the number of basic solutions required, the author wishes to point out that five independent basic solutions are required for the case of flow through a blade row with one splitter vane if one desires to construct all possible flows through the blade row. The five physical conditions that may be varied by appropriate combinations of these solutions are (1) the location of the rear stagnation point on the blade, (2) the location of the rear stagnation point on the splitter vane, (3) the weight flow, (4) the rotational speed, and (5) the amount of prewhirl. If it were desired to obtain the solution for nonzero prewhirl, the appropriate substitution in Equation [11] for $w_{\theta,1}$ would have to be made in determining the linear relation among A_0, A_1, A_2, A_3, and A_4 defined by that equation. In the set of four basic solutions suggested by Mr. Ellis, the amount of prewhirl cannot be changed from the initial value (zero in the example which he shows). However, the solutions are adequate for the construction of all possible flows with zero prewhirl, that is, for any rotational speed or flow rate. As he suggests at the end of his comments, a fifth solution would be required in order to construct solutions with varying amounts of prewhirl.

As to the values suggested for boundary values in Table 1, the author chose those values solely for the purpose of illustrating the independence of the basic solutions. In the actual numerical solution of the problem the use of boundary values, such as Mr. Ellis suggests, should certainly result in some saving in computing time.

In regard to Mr. Hamrick's inquiry as to the time required for the solution of flow in a blade row with a splitter vane, the author is unable to give an estimate of time required based on experience. The time required for such a solution would be a function of the kind of computing equipment available and the experience of the programmer. With advanced type computing equipment the actual computing time should be quite small, of the order of an hour. The time-consuming part of the job would be the programming time and research necessary to decide upon an appropriate method of solving the finite difference equations. In view of these considerations it is impossible to estimate the time required for one solution. However, one can say that succeeding solutions could be obtained in a much shorter time.

Two-Phase Flow in Rough Tubes

By D. CHISHOLM[1] AND A. D. K. LAIRD[2]

This paper presents data for pressure drop and saturation during flow of air-water mixtures in smooth and rough horizontal tubes. Improvements in the two-phase flow correlations for rough tubes are presented. Approximate empirical relationships developed using these improvements correlated the majority of the data within 15 per cent.

Nomenclature

The following nomenclature is used in the paper:

A_L = cross section occupied by liquid, sq ft
A_P = cross section of tube, sq ft
C = constant in Equation [7]
C' = constant in Equation [5]
D = tube diameter
G_G = gas-mass velocity based on tube cross section: $\rho_G V_{GP}$, lb/sec (ft²)
G_L = liquid-mass velocity based on tube cross section = $\rho_L V_{LP}$, lb/sec (ft²)
g = gravitational acceleration, ft/sec²
ΔL = an increment of distance in direction of motion
m = exponent of $\bar{X}$ in Equation [7]
n = exponent of N_R in Equation [5]
N_R = Reynolds number
N_{RLP} = Reynolds number where the liquid flows alone = $G_L D/\mu_L$
N_{RL} = Reynolds number based on actual liquid velocity during two-phase flow = $G_L D/\mu_L R_L$
P = mean system pressure, psf
ΔP_{TP} = total (friction) pressure drop over increment of length for both phases flowing simultaneously, psf
ΔP_G = friction pressure drop for gas flowing alone in tube, psf
ΔP_L = friction pressure drop for liquid flowing alone in tube, psf
ΔP_M = momentum pressure drop for two-phase flow, psf
R_L = liquid saturation = A_L/A_P
V_G = actual mean gas velocity during two-phase flow, fps
V_{GP} = mean gas velocity as if gas flows alone in tube, fps
V_L = actual mean liquid velocity during two-phase flow, fps
V_{LP} = mean liquid velocity as if liquid flows alone in tube, fps
X = Martinelli parameter $\sqrt{(\Delta P_L/\Delta P_G)}$
$\bar{X}$ = modified Martinelli parameter defined by Equation [6b]
μ_L = absolute viscosity of liquid, lb/sec ft
ρ_L = liquid density, pcf
ρ_G = gas density, pcf
λ = pipe friction factor for rough tube
λ' = pipe friction factor in general
λ_s = pipe friction factor for smooth tube
ϵ/D = pipe roughness ratio

[1] English Electric Company, Harwell, England; formerly, Research Fellow, University of California, Berkeley, Calif.

[2] Associate Professor of Mechanical Engineering, University of California, Berkeley, Calif.

Contributed by the Fluid Mechanics Subcommittee of the Hydraulics Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 1, 1956. Paper No. 57—SA-11.

Introduction

The flow of gas-liquid and vapor-liquid mixtures, or, as they are commonly called, two-phase mixtures, occurs frequently in industry. It is only within the past two decades, however, that this subject has been investigated scientifically. The most extensive investigations were carried out at the University of California and led to the well-known Lockhart-Martinelli correlations (1, 2, 3).[3] These correlations, although based on smooth-tube data, have been found useful for many diverse conditions, including flow in rough tubes. However, the greater the deviation of a set of conditions from flow in a smooth tube, the greater is the error in prediction.

The present investigation was initiated to discover correlations more satisfactory for horizontal rough-tube conditions. Pressure drop and saturation data were obtained for the flow of air-water mixtures in a smooth tube and several rough tubes, all of approximately 1-in. bore. The tubes were in the horizontal plane, and pressures were close to atmospheric throughout.

Equipment and Procedure

A diagrammatic sketch of the equipment is shown in Fig. 1. The test length consisted of an 8-ft tube length of approximately 1-in. bore, with pressure taps at 2-ft intervals over this length. The pressure taps were connected to water-air manometers; water filled the pressure lines from tap to manometer. Before a manometer reading was taken, water was passed through the pressure lines to insure that they contained no air. The saturation was measured by trapping the liquid in the test section by simultaneously shutting "quick-closing" cocks at entry and exit, then measuring the trapped liquid. This method was first developed by Moore and Wilde (5) some 20 years ago. The two quick-closing cocks were approximately 12 ft apart, 2 ft from each end of the test length, and were interconnected by a linkage system to permit their simultaneous closure. A small cock situated in the body of the lower quick-closing cock enabled the trapped water to be drained from the test section. To facilitate drainage, the tube was inclined 7 deg from the horizontal and air was blown through it. Calibration tests were carried out to determine

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

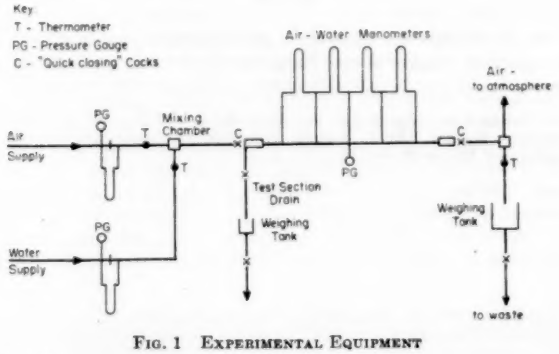


FIG. 1 EXPERIMENTAL EQUIPMENT

the quantity of water left behind on the tube wall during draining by subtracting the volume recovered by the foregoing drainage method from the volume required to fill the dry tube.

The mixing chamber was 7 ft upstream from the first pressure-tapping point, and the separator 4 ft downstream from the last pressure-tapping point. The separator and mixing chamber were therefore 19 ft apart. The mixing chamber consisted essentially of the main tube and a smaller concentric tube within it; the air was admitted through the smaller tube and the water through the annulus between the tubes. The gravity separator discharged the air to atmosphere and the water to a measuring tank.

The air and water flow rates were measured by orifices and manometers; the water also was measured gravimetrically. The mean system pressure was measured at the center tapping point by water manometer and Bourdon pressure gage at low and high pressures, respectively. Temperatures at entry to the mixer and exit from the separator were measured by thermometers.

Two viewing sections 6 in. long, located 1 ft upstream and downstream, respectively, from the test length permitted observation of the flow patterns. These viewing sections were machined from Lucite blocks with parallel outside walls to reduce refraction.

Tube Surfaces. The tube surfaces tested were as follows:

(*a*) Smooth brass tube. Bore 1.062 in.; ϵ/D:0.000.

(*b*) Commercial galvanized tube. Bore 1.043 in.; ϵ/D:0.0025.

(*c*) Brass tube with concrete internal surface. The concrete was applied to a brass tube with an internal thread (tube *d*). An irregular finish was obtained. Bore 1.059 in.; ϵ/D:0.013.

(*d*) Brass tube with internal thread (see Fig. 2). Bore 1.077 in. Measured $\epsilon/D = 0.028$; apparent ϵ/D:0.037.

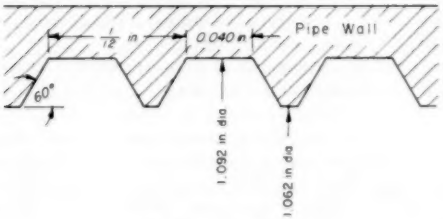


FIG. 2 TUBE THREAD

(*e*) Sand distributed nonuniformly on brass tube. Bore 1.032 in.; ϵ/D:0.045.

(*f*) Sand distributed uniformly on galvanized tube. Bore 1.018 in.; ϵ/D:0.068.

Shellac was used to glue the sand to the tube walls. With the exception of the threaded tube no direct measurements of the surface roughnesses were made. The values quoted are values obtained from a Moody chart (6), extrapolated where necessary, corresponding to the measured friction factor in the region of complete turbulence.

Experimental Data. The ranges of water and air-mass velocities were from 39 to 600 lb/sec ft^2 and 0.1 to 20 lb/sec ft^2, respectively, corresponding approximately to a range of Reynolds number from 4500 to 80,000 for the water, and 1000 to 140,000 for the air. Mass velocities, throughout the paper, were evaluated using the tube cross-sectional area, and both phases were considered turbulent throughout their ranges. Temperatures were in the range from 59 to 74 F, corresponding to a water-viscosity range from 77 $\times 10^{-6}$ lb/ft sec to 63 $\times 10^{-6}$ lb/ft sec, and an air-viscosity range from 1.20 $\times 10^{-6}$ lb/ft sec to 1.23 $\times 10^{-6}$ lb/ft sec. The data are on file with the American Document Institute.[4]

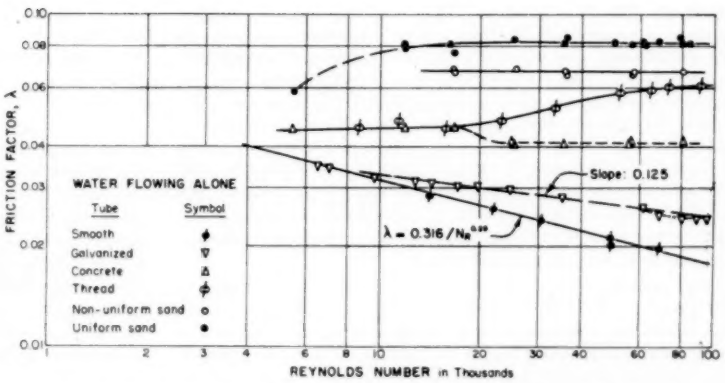


FIG. 3 TUBE FRICTION FACTORS AS A FUNCTION OF REYNOLDS NUMBER

The tube friction factors obtained from measured pressure drops during water flow are shown in Fig. 3.

Mean arithmetic values of the water temperature at the mixing chamber and separator, respectively, were used throughout the paper in determining the water properties. In determining the air properties, no appreciable error was committed by assuming the air was dry and at the mean temperature of the water.

THE PHENOMENA CONSIDERED

Several types of flow are encountered in the simultaneous concurrent passage of a liquid and a gas through a horizontal tube as the flow rate of each is increased from zero (1). When the liquid and the gas rates are both small, the flow closely resembles that in an open channel and is called separated flow. If the water rate is then held approximately constant and the gas rate is increased, the liquid surface develops waves. Secondary currents set up by the velocity gradients in the liquid may cause the liquid to climb the tube walls and, aided by the waves, to coalesce at the top. Further increase of the gas rate may result in annular flow which consists of a continuous core of gas surrounded by a layer of liquid adhering to the tube walls. More rapid gas rate may cause the liquid walls to disintegrate into drops carried along in the gas inside wetted tube walls. This type is called mist flow. If in the previously mentioned separated flow, more liquid were flowing before the gas velocity was increased, the liquid might fill the tube for short lengths separated by large bubbles of gas. In most cases many small bubbles would be mixed in the liquid. Such a flow is often called slug flow. Still larger gas rates would again cause annular flow and eventual breakdown as in the previous case. If the liquid rates were to be increased at the same time as the gas rate was increased, the flow would probably consist of gas bubbles fairly uniformly dispersed throughout the liquid.

At present the type of flow to be expected with given flow rates, and the conditions at which transitions between types of flow will

[4] The data have been deposited as Document number 5178 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting $2.50 for photoprints, or $1.75 for 35 mm microfilm. Advance payment is required. Make checks or money order payable to: Chief, Photoduplication Service, Library of Congress.

occur, are difficult to predict. The type of flow depends upon the amount of the two fluids present and their relative velocities. In the bubble and slug types, the gas and liquid flow at approximately the same velocities with the gas always slightly faster because it tends to stay near the center line of the tube where the velocity is highest. When the gas rate is high enough relative to the liquid rate, the gas forms a continuous core which may contain a large portion of the liquid. When a continuous gas core is formed the gas velocity is much higher than the liquid velocity. In horizontal flow there is no back flow except at the small liquid rates when more of the liquid may be carried forward by the gas than is allowed to flow out of the tube.

The amounts of the two components present are expressed as saturations. The saturation of a component is calculated as the volume of the component in a length of tube, divided by the total volume in the same length of the tube. The velocity of each component is calculated as if it alone flowed in the tube at the same volume rate.

If the Reynolds number based on this definition of velocity and tube diameter is over 5000 the flow is turbulent; if less than 1000 it is laminar. It is possible to have either of the components in laminar flow and the other in turbulent flow, or both laminar, or both turbulent. The present research covers much of the range of the bubble, slug, and annular types of flow, but not the separated or the mist types. In all cases both the gas and liquid flows were turbulent.

Bases of Correlation. The usual methods of correlating two-phase-flow pressure-drop data use the pipe-friction-factor variation with Reynolds number and wall roughness for homogeneous single-phase flow in tubes. The apparent reason for the success of all such correlations is that the Reynolds number and the roughness have a significant effect upon two-phase flow. Since one of the components wets the tube, the roughness has its effect on this component and thereby the roughness effect enters the system. Reynolds number plays its usual role in both fluids by specifying in dimensionless terms the absolute and the relative velocities of the fluids and their approximate turbulence level. The saturation of the two components can be correlated by the same parameters. The saturation is of interest for its influence on heat transfer and actual fluid velocities, particularly during phase change. The correlations for saturation have not yet become sufficiently accurate for use in predicting pressure drop. They are of use in estimating the relative quantities of the two fluids present to check the results of analyses based on hydrodynamic theory.

It appears incontestable that the Reynolds number, the tube roughness, and the flow rates uniquely determine the characteristics of isothermal two-phase flow. To date it has been impossible to solve the general hydrodynamic two-phase-flow problem. In view of the fact that the single-phase problem is not completely solved, it seems likely that the two-phase problem will remain unsolved for many years. Consequently, the next best treatment of the problem is a correlation of the data based on the influence of the known parameters. Since the friction factor is effective in single-phase flow, it should appear in the correlation. There remains the choice of the form of the correlating parameters. The Martinelli parameters were devised to correlate a small range of smooth-tube data in the annular-flow regime. It has been found, however, that his correlation method is applicable far beyond the range for which it was proposed. In fact, it is difficult to postulate a two-phase-flow system for which his parameters are ineffective. Consequently, they must contain some fundamental principles.

There are other parameters consisting of groups of the same variables as used in Martinelli's parameters which are valid for correlating certain data well, or for all two-phase data, more or less well. In general, the Martinelli correlation method has been the most universally applicable. Even it, however, cannot be used directly for the accurate prediction of rough-tube data. To preserve the universal nature of the Martinelli correlation, and at the same time extend its range of applicability, the pressure-drop parameter was not changed, but the flow-rate-ratio parameter X was altered. The new parameter $\bar{X}$ reduces to X for smooth tubes and retains that part of Martinelli's X which is considered to be its essence for application to rough-tube data. Further, the form of the empirical functional relations between the pressure drop and X, and the saturation and X, were changed to a series of decreasing powers of $\bar{X}$. This choice was made, not only because it best fits the present data, but because this form has been used in correlating data from two-phase fluid-solid systems and two-phase fluid systems with heat and mass transfer. In adapting the new parameter $\bar{X}$, it has been tacitly assumed that the larger the range of applicability of a correlating method, the more fundamental is its basis.

Correlation of Data

Pressure Drops. The Lockhart-Martinelli correlation for smooth tubes gave a single curve for turbulent-turbulent flow on plotting $\sqrt{(\Delta P_{TP}/\Delta P_L)}$ to a base of X equal to $\sqrt{(\Delta P_L/\Delta P_G)}$. Inspection of this curve indicates that it lies within ±30 per cent of the equation

$$\frac{\Delta P_{TP}}{\Delta P_L} = 1 + 21/X + 1/X^2 \quad [1]$$

This also may be expressed

$$\Delta P_{TP} = \Delta P_L + 21\sqrt{(\Delta P_L \Delta P_G)} + \Delta P_G \quad [2]$$

The present smooth-tube data may be correlated satisfactorily for X greater than 0.4 by the reduced form of Equation [1]

$$\frac{\Delta P_{TP}}{\Delta P_L} = 1 + 21/X \quad [3]$$

Hence as will be seen in Fig. 4 a logarithmic plot of $(\Delta P_{TP}/\Delta P_L) - 1$ to a base of $\bar{X}$ will give a straight line for values of $\bar{X}$ above 0.4. As the great majority of the data were in the region of $\bar{X}$ greater than 0.4, this system of ordinates was selected in correlating the rough-tube pressure-drop data.

The Martinelli parameter X may be expressed in terms of the various flow and physical properties as

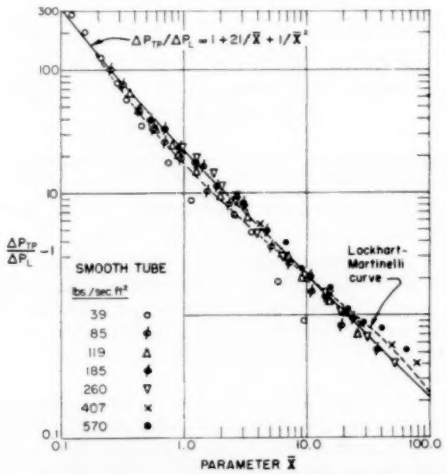


Fig. 4 $(\Delta P_{TP}/\Delta P_L) - 1$ as a Function of $\bar{X}$ for Smooth Tube

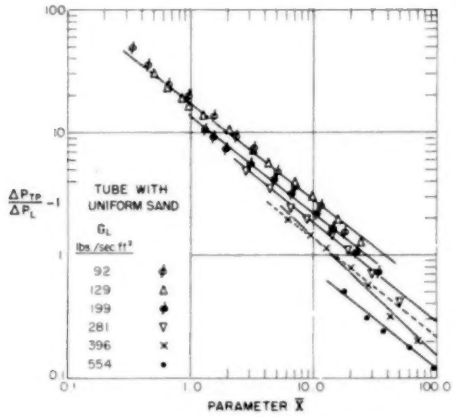


FIG. 5 $(\Delta P_{TP}/\Delta P_L) - 1$ AS A FUNCTION OF X FOR TUBE WITH UNIFORM SAND ROUGHNESS

$$X = \left(\frac{G_L}{G_G}\right)^{\frac{2-n}{2}} \left(\frac{\mu_L}{\mu_G}\right)^{\frac{n}{2}} \left(\frac{\rho_G}{\rho_L}\right)^{0.5} \quad \ldots\ldots\ldots [4]$$

where n is the power of Reynolds number in the friction-factor relation

$$\lambda' = C' N_R^{-n} \quad \ldots\ldots\ldots [5]$$

For smooth tubes, the Blasius equation gives $n = 0.25$, hence

$$X = \left(\frac{G_L}{G_G}\right)^{0.875} \left(\frac{\mu_L}{\mu_G}\right)^{0.125} \left(\frac{\rho_G}{\rho_L}\right)^{0.5} \quad \ldots\ldots\ldots [6a]$$

It is important to appreciate that, while the work of Lockhart and Martinelli has indicated that Equation [6*a*] provides a satisfactory parameter for the correlation of two-phase-flow smooth-tube data, their investigations have not confirmed that Equation [4] provides a satisfactory parameter for the correlation of rough-tube data; nevertheless they have not explicitly excluded this by defining X as $\sqrt{(\Delta P_L/\Delta P_G)}$, rather than as Equation [6*a*]. With rough tubes the value of n approaches zero; consequently Equation [4] suggests that the pressure drop should tend to become a function of G_L/G_G to the power unity. The present investigation has not found this to be the case and, hence, for this reason and the considerations mentioned previously, the general form of X given in Equation [4] is not considered a suitable parameter for correlating two-phase-flow data. The present investigation has indicated that the most satisfactory correlations are obtained using X calculated, regardless of the value of n, by Equation [6*a*]. This is tantamount to redefining the parameter for X as defined by Lockhart and Martinelli; consequently the symbol $\bar{X}$ is introduced where

$$\bar{X} \equiv \left(\frac{G_L}{G_G}\right)^{0.875} \left(\frac{\mu_L}{\mu_G}\right)^{0.125} \left(\frac{\rho_G}{\rho_L}\right)^{0.5} \quad \ldots\ldots\ldots [6b]$$

For the special case of $n = 0.25$, $\bar{X}$ is, of course, identical to X as will be seen by comparing Equations [6*a*] and [6*b*].

Pressure-drop data correlated using $\bar{X}$ are shown in Fig. 5 for the tube with uniform sand roughness. It will be observed that the data, as anticipated, fall close to lines of the form

$$\frac{\Delta P_{TP}}{\Delta P_L} = 1 + C/\bar{X}^m \quad \ldots\ldots\ldots [7]$$

where C and m are constants for a particular liquid-flow rate and tube surface. The values of C and m, obtained by applying the "method of least squares" to the logarithmic plot of $(\Delta P_{TP}/\Delta P_L) - 1$ to $\bar{X}$, are given in Table 1. Only a few tests had $\bar{X}$ values less than 0.4 and these tests have been excluded from the present analysis.

It was not possible to obtain satisfactory correlations for the tests at the lowest liquid rate, and no values are shown in Fig. 5 for this flow rate, nor are values of C and m given in Table 1. As the Reynolds number at the lowest liquid-flow rate is approximately 5000, the difficulty in obtaining satisfactory correlations is undoubtedly due to the transition from laminar to turbulent flow.

Equation [7] may be expressed

$$\frac{\Delta P_{TP}}{\Delta P_L} = 1 + C \left(\frac{G_G}{G_L}\right)^{0.875m} \left(\frac{\mu_G}{\mu_L}\right)^{0.125m} \left(\frac{\rho_L}{\rho_G}\right)^{0.5m} \quad \ldots [8]$$

It must be stressed that the limited range of viscosities and densities used in the present tests precludes a definite confirmation of the powers of the viscosity and density ratios in this equation.

Only approximate correlations for C and m have been developed so far. In Figs. 6 and 7, C and m are shown to be functions of $(\lambda/\lambda_S)\sqrt{N_{RLP}}$ and λ/λ_S, respectively, where λ is the friction factor for a rough tube, λ_S the friction factor for a smooth tube at the same liquid Reynolds number N_{RLP} based on the pipe diameter. Figs. 6 and 7 may be used to predict pressure drops in rough tubes in the following manner:

1 Reynolds number (N_{RLP}) is calculated and the friction factors (λ) for the rough tube estimated in the conventional manner for homogeneous flow. The corresponding friction factor (λ_S) for a smooth tube also is estimated.

2 The liquid friction pressure drop (ΔP_L) is evaluated on the assumption of the liquid alone in the rough tube.

3 The ratio λ/λ_S is evaluated, and m obtained from Fig. 7.

4 The term $(\lambda/\lambda_S)\sqrt{N_{RLP}}$ is evaluated, and C obtained from Fig. 6.

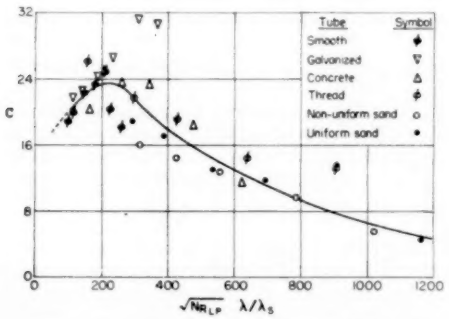


FIG. 6 C AS A FUNCTION OF $(\lambda/\lambda_S)\sqrt{(N_{RLP})}$

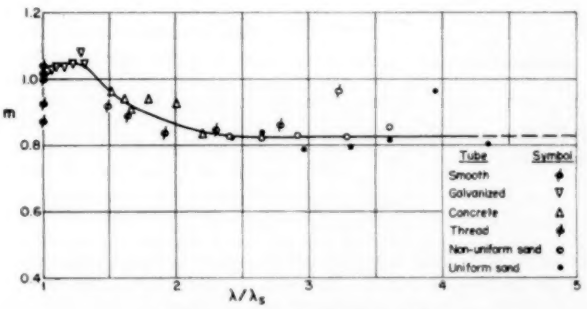


FIG. 7 m AS A FUNCTION OF λ/λ_S

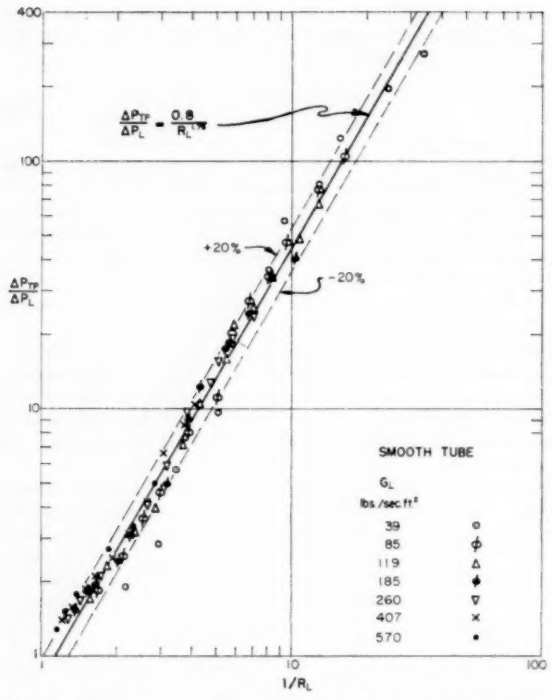


FIG. 8 $\Delta P_{TP}/\Delta P_L$ AS A FUNCTION OF $1/R_L$ FOR SMOOTH TUBE

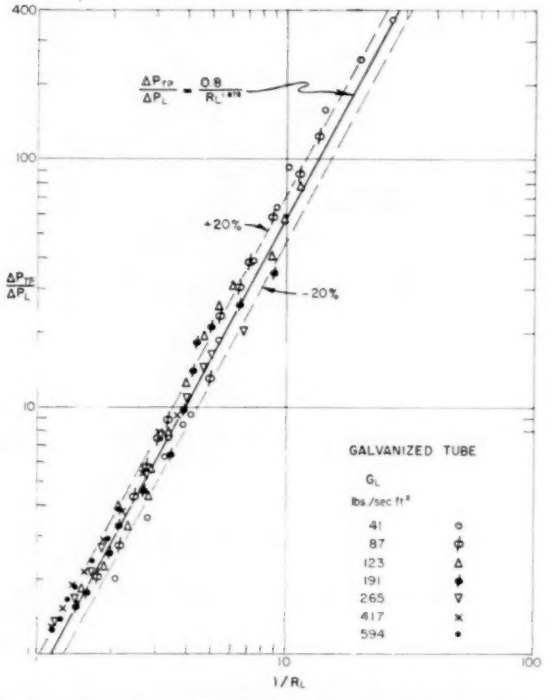


FIG. 9 $\Delta P_{TP}/\Delta P_L$ AS A FUNCTION OF $1/R_L$ FOR GALVANIZED TUBE

5 Using the foregoing values of ΔP_L, m, and C, the two-phase pressure drop is calculated from Equation [7].

No correction has been made in this analysis for momentum effects. In Appendix 1, a maximum momentum pressure drop of 8.5 per cent is estimated. This value is the result of the assumption that the gas and liquid have the same percentage increase in velocity over finite lengths of pipe. Because of the difference in densities of the phases, this value must be too high, possibly by a considerable amount. The majority of the data must have momentum pressure drops of the order of 1 or 2 per cent of the total pressure drop.

Saturation—Pressure Drop Correlation. The two-phase pressure-drop ratio is shown as a function of the reciprocal of the liquid saturation for the smooth tube in Fig. 8, the galvanized tube in Fig. 9, and the uniform sand-roughness tube in Fig. 10. It will be observed that the majority of the data for the smooth and the galvanized tubes lie within ±20 per cent of the equations:

Smooth tube

$$\Delta P_{TP} = 0.8\Delta P_L/R_L^{1.75} \quad \ldots\ldots [9]$$

Galvanized tube

$$\Delta P_{TP} = 0.8\Delta P_L/R_L^{1.875} \quad \ldots\ldots [10]$$

It is shown in Appendix 2 that both Equations [9] and [10] can be reduced to the form

$$\Delta P_{TP} = 0.8\,\frac{\lambda'\Delta L V_L^2 \rho_L}{2gD} \quad \ldots\ldots [11]$$

where V_L is the mean liquid velocity during two-phase flow based on the liquid cross section. However, as will be seen in Fig. 10, with greater surface roughness the data no longer lie on a single line, and no simple relationships between liquid velocity and

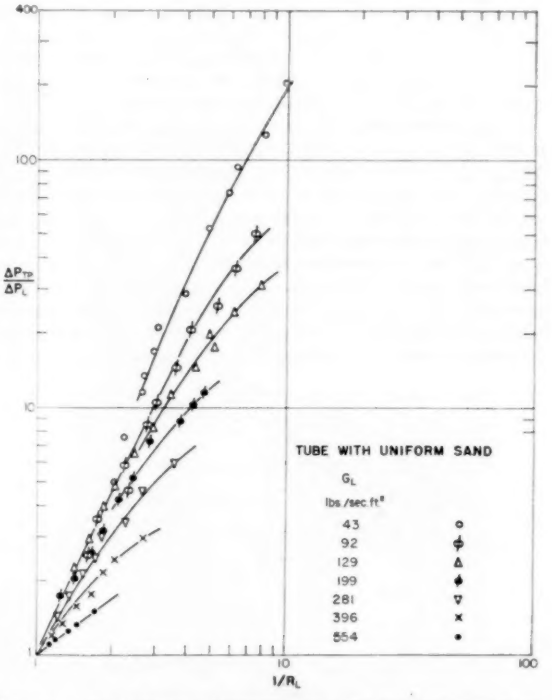


FIG. 10 $\Delta P_{TP}/\Delta P_L$ AS A FUNCTION OF $1/R_L$ FOR UNIFORM SAND-ROUGHNESS TUBE

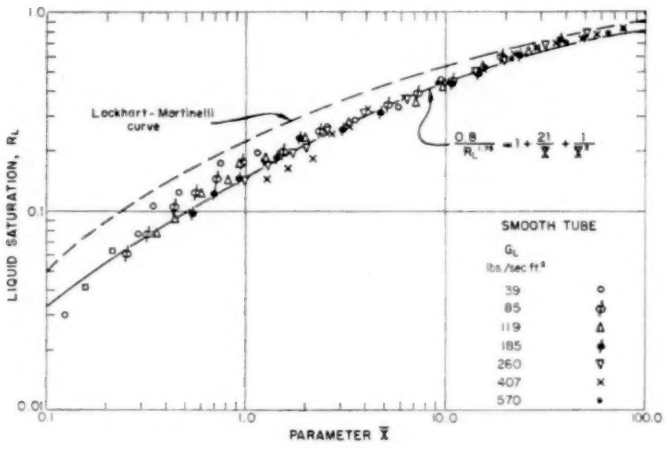


FIG. 11 R_L AS A FUNCTION OF $\bar{X}$ FOR SMOOTH TUBE

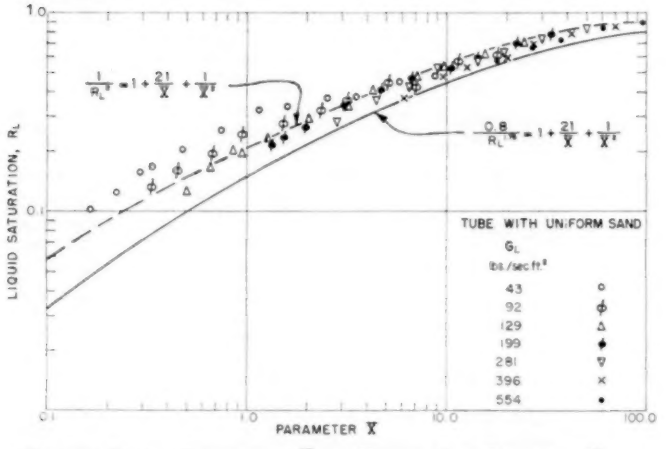


FIG. 12 R_L AS A FUNCTION OF $\bar{X}$ FOR UNIFORM SAND-ROUGHNESS TUBE

pressure drop such as Equation [11] are obtained. The data for the threaded tube show the same trends as for the uniform sand-roughness tube.

Saturation Correlations. Lockhart and Martinelli correlated the liquid saturation with the parameter X. The saturation data plotted as a function of $\bar{X}$ are shown in Fig. 11 for the smooth tube and in Fig. 12 for the tube with uniform sand roughness. As $\bar{X}$ is identical to the Martinelli parameter X for smooth tubes, the mean curve through the data in Fig. 11 can be obtained by equating Equations [1] and [9]

$$0.8/R_L^{1.75} = 1 + 21/X + 1/X^2 = 1 + 21/\bar{X} + 1/\bar{X}^2 \quad [12]$$

This curve is shown in Fig. 11. With the exception of the tests for the smallest liquid rate, the data can be seen to fall within ±25 per cent of this equation. The equation for the galvanized tube, which correlates the data with similar accuracy, is

$$0.8/R_L^{1.875} = 1 + 26/\bar{X} \quad [13]$$

where 26 is the mean value of C in Table 1. The satisfactory correlation obtained with this equation in the absence of higher powers of $\bar{X}$ is due, presumably, to the restricted range of $\bar{X}$-values.

These equations suggest a similar form for the rougher tubes, and the equation

$$1/R_L^2 = 1 + 21/\bar{X} + 1/\bar{X}^2 \quad [14]$$

can be seen in Fig. 12 to correlate the data for the tube with the uniform sand within ±25 per cent, again with the exception of the lowest liquid-flow rate. An equation of similar accuracy for the threaded tube is

$$0.9/R_L^2 = 1 + 21/\bar{X} + 1/\bar{X}^2 \quad [15]$$

In Figs. 11 and 12 it will be observed that for $\bar{X}$ less than 3.5 there is a noticeable trend with liquid-flow rate; the saturation for a particular $\bar{X}$-value increases with decreasing liquid-flow rate. It has not as yet been possible to develop satisfactory correlations for this phenomenon. For $\bar{X}$-values greater than 3.5, data at all flow rates tend to fall on a single curve.

DISCUSSION

The procedure adopted here, of plotting $(\Delta_{TP}/\Delta P_L) - 1$ to $\bar{X}$, affords many advantages over the plot of $\sqrt{(\Delta P_{TP}/\Delta P_L)}$ to X developed by Lockhart and Martinelli, including ease of computation and accuracy of prediction. It constitutes a basic improvement in two-phase flow correlations. The logarithmic plots readily indicate that the pressure drop may be evaluated by equations of the form

$$\Delta P_{TP}/\Delta P_L = 1 + C/\bar{X}^m \quad [7]$$

The present investigation has shown that this form of equation applies for both smooth and rough tubes. It is expected that equations of this form will be found to hold for a wide range of conditions. Lockhart and Martinelli's data for air and several liquids for $\bar{X}$ greater than 0.4 may be correlated satisfactorily using Equation [7]. Also, it should be noted that Equation [7] is similar in form to equations obtained with solids-gas flow, where logarithmic plots of $(\Delta P_{TP}/\Delta P_L) - 1$ to a base of a function of the flow and physical properties also give linear relationships (7).

Undoubtedly, for $\bar{X}$-values less than 0.4, more terms in $\bar{X}$ would have to be added to Equation [7]. Insufficient data were obtained with the present investigation to enable the determination of these powers, although Lockhart and Martinelli's data suggested the powers given in Equation [1]. The further investigation that is required in the region of low $\bar{X}$-values will be complicated by the fact that in this region the pressure drop will be considerably greater than in the present tests, and, in consequence, the momentum forces will be of such a magnitude that they may no longer be neglected satisfactorily.

The correlations of C and m given here, while approximate, enable one to predict the pressure drop for the majority of the tests within ±15 per cent of the experimental values. The accuracy of prediction is illustrated by the plot in Fig. 13 of the estimated pressure-drop ratio to the measured pressure-drop ratio for the tube with uniform sand. The greatest deviations (−25 per cent) occurred with the galvanized tube. Fig. 14 illustrates the correlation obtained with this tube. Direct application of the smooth-tube formula (Equation [3]) gives the less satisfactory maximum deviation of −46 per cent.

The difficulty in obtaining accurate values for C and m may be illustrated by the data from the tube with uniform sand at the

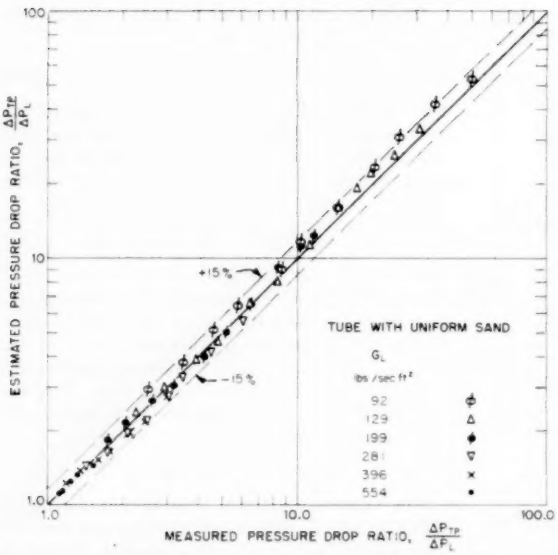


FIG. 13 ESTIMATED PRESSURE DROP VERSUS MEASURED PRESSURE DROP FOR UNIFORM SAND ROUGHNESS TUBE

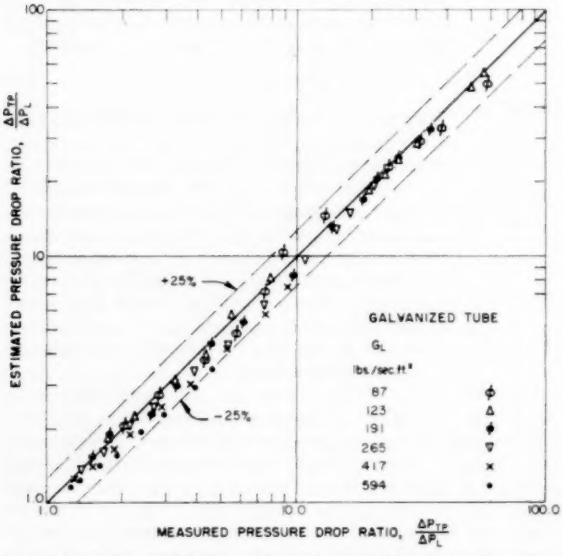


FIG. 14 ESTIMATED PRESSURE DROP VERSUS MEASURED PRESSURE DROP FOR GALVANIZED TUBE

liquid rate of 396 lb/sec ft². In Fig. 5, the two points at the largest values of $\bar{X}$ have undue weight in the least-squares method of curve fitting. The broken line might be at a better slope to give a value of m more consistent with the other values of m given in Table 1 for this tube. One might justify this treatment by noting that unity must be subtracted from the pressure-drop-ratio measurement, which is itself close to unity. However, the total number of points for this curve is too small to justify disregarding the two low points, especially since the other curves of Fig. 5 also tend to drop below the straight lines at higher values of $\bar{X}$.

The form of the equations for the rough tubes may be similar to those obtained during steam-water flow with evaporation. The bubbles forming at the wall may produce a disruption of the laminar layer at the wall in a similar manner to that produced by surface roughness. Limited confirmation of this concept is given by the data of Stein, et al. (4) for steam-water mixtures flowing downward in an annulus with evaporation, where the data fall close to the curve

TABLE 1 SUMMARY OF PRESSURE DROP CORRELATIONS

Smooth Tube		Dia: 1.062"		ε/D ±0.000		
G_L lb/sec. (sq.ft)	85	119	185	260	407	570
Mean Reynolds Number, N_{RLP}	10,000	14,100	21,800	30,700	48,100	67,300
λ/λ_S	1.0	1.0	1.0	1.0	1.0	1.0
m	1.04	1.0	1.02	1.02	0.927	0.875
C	18.7	19.9	22.1	23.4	20.2	18.0
Maximum* Percentage Deviation	+15/-12	+5/-4	+11/-10	+8/-8	+14/-12	+7/-4
Galvanized Tube		Dia: 1.043"		ε/D ±0.0025		
G_L lb/sec.(sq.ft)	87	123	191	265	417	594
Mean Reynolds Number, N_{RLP}	11,800	16,500	25,900	35,900	56,600	79,400
λ/λ_S	1.07	1.09	1.16	1.22	1.29	1.31
m	1.03	1.04	1.04	1.05	1.08	1.05
C	21.7	22.6	24.3	26.6	31.4	30.8
Maximum* Percentage Deviation	+19/-9	+6/-6	+10/-10	+6/-9	+5/-4	+3/-3
Tube with Concrete Surface		Dia: 1.059"		ε/D ±0.013		
G_L lb/sec.(sq.ft)	84	118	184	259	404	575
Mean Reynolds Number, N_{RLP}	11,700	16,300	25,600	36,000	56,000	80,000
λ/λ_S	1.51	1.62	1.65	1.79	2.0	2.2
m	0.96	0.94	0.90	0.94	0.92	0.83
C	20.3	25.0	23.6	23.3	18.4	11.5
Maximum* Percentage Deviation	+7/-9	+8/-10	+7/-7	+5/-5	+6/-6	+3/-4
Tube with Internal Thread		Dia: 1.077"		ε/D ±0.037		
G_L lb/sec.(sq.ft)	82	114	176	251	392	555
Mean Reynolds Number, N_{RLP}	11,400	15,900	23,900	34,600	53,200	78,600
λ/λ_S	1.49	1.63	1.92	2.30	2.79	3.22
m	0.92	0.88	0.83	0.84	0.86	0.96
C	26.1	24.9	21.8	19.1	14.4	13.2
Maximum* Percentage Deviation	+19/-10	+9/-12	+12/-14	+5/-5	+10/-7	+3/-3
Tube with Non-Uniform Sand		Dia: 1.032"		ε/D ±0.045		
G_L lb/sec.(sq.ft)	- -	124	194	268	426	604
Mean Reynolds Number, N_{RLP}	- -	16,800	26,100	36,300	57,000	80,600
λ/λ_S	- -	2.41	2.64	2.92	3.28	3.6
m	- -	0.83	0.82	0.83	0.83	0.85
C	- -	16.0	14.5	12.9	9.7	5.7
Maximum* Percentage Deviation	- -	+14/-8	+5/-4	+2/-3	+3/-2	+1/-2
Tube with Uniform Sand		Dia: 1.018"		ε/D ±0.068		
G_L lb/sec.(sq.ft)	92	129	199	281	396	554
Mean Reynolds Number, N_{RLP}	12,300	17,000	26,300	37,100	52,700	73,200
λ/λ_S	2.65	2.96	3.31	3.60	3.94	4.30
m	0.84	0.79	0.80	0.81	0.96	0.80
C	18.9	17.1	13.1	11.9	12.6	4.65
Maximum* Percentage Deviation	+14/-8	+9/-7	+4/-8	+4/-7	+6/-5	+1/-3

*Percentage deviation in predicted pressure drop using Equation [7] and tabulated value of C and m.

$$\Delta P_{TP}/\Delta P_L = 1 + 30/\bar{X}^{0.8} \quad [16]$$

The power of $\bar{X}$ lies close to the values obtained with the very rough tubes used in the present investigation; the value of the constant C, however, is greater, probably because Stein's liquid-pressure drops were evaluated from smooth-tube data.

Further study is required to find the influence of tube diameter and the system pressure. The analysis of Martinelli and Nelson (8) with steam-water mixtures at pressures above 500 psi, under conditions when the gas properties approach the liquid properties, suggests that an increase in pressure will be associated with a decrease in the value of C in Equation [7]. Examination of the data of Thomsen and Ravenscroft (9, 10) for the turbulent flow of benzene-air and water-air, respectively, in smooth $^1/_2$-in-diam tubes at pressures of 50 psia, indicates that the two-phase pressure drop is proportional to G_L/G_G to the power 1.3, as compared with the power 0.875 found for 1-in. tubes

Conclusions

1 For correlating two-phase flow data, logarithmic plots of $(\Delta P_{TP}/\Delta P_L) - 1$ as a function of the parameter $\bar{X}$ are recommended, in preference to the plot of $\sqrt{(\Delta P_{TP}/\Delta P_L)}$ as a function of X recommended by Lockhart and Martinelli.

2 The data of Lockhart and Martinelli for the turbulent-turbulent flow of air and any of a number of liquids may be satisfactorily correlated by Equation [1]. For $\bar{X}$-values greater than 0.4, Equation [3], which is a reduced form of Equation [1], gives satisfactory prediction of the pressure drop for the majority of the writers' tests with the smooth tube, within ±20 per cent of the experimental value. These equations bear considerable resemblance to the form of equations obtained with solids-gas flow.

3 Equations of the form of Equation [7] are obtained with rough tubes. Both C and m decrease with increasing surface roughness. Approximate correlations for C and m have been found, and their use gives pressure-drop predictions within ±15 per cent of the experimental values for the majority of the tests. The maximum deviation to be expected is ±25 per cent.

4 Equation [11] correlates the majority of the smooth and galvanized-tube data within ±20 per cent. No satisfactory correlations of this form could be obtained with the rougher tubes.

5 Equations for the prediction of liquid saturation have been developed. They give values within ±25 per cent of the experimental results for the majority of the tests.

BIBLIOGRAPHY

1 "Isothermal Pressure Drop for Two-Phase, Two-Component Flow in a Horizontal Pipe," by R. C. Martinelli, L. M. K. Boelter, T. H. M. Taylor, E. G. Thomsen, and E. H. Morrin, Trans. ASME, vol. 66, 1944, pp. 139–151.

2 "Two-Phase, Two-Component Flow in the Viscous Region," by R. C. Martinelli, J. A. Putnam, and R. W. Lockhart, Trans. AIChE, vol. 4, 1946, pp. 681–705.

3 "Proposed Correlation of Data for Isothermal Two-Phase, Two-Component Flow in Pipes," by R. W. Lockhart and R. C. Martinelli, *Chemical Engineering Progress*, vol. 45, 1949, pp. 39–48.

4 "Pressure Drop and Heat Transfer to Nonboiling and Boiling Water in Turbulent Flow in an Internally Heated Annulus," by R. P. Stein, J. W. Hoopes, Jr., M. Markels, Jr., W. A. Selke, A. J. Bendler, and C. F. Bonilla, American Institute of Chemical Engineers, Chemical Engineering Progress Symposium Series No. 11, vol. 50, 1950.

5 "Experimental Measurement of Slippage in Flow Through Vertical Pipes," by T. V. Moore and H. D. Wilde, Jr., Trans. AIME, Petroleum Division, vol. 92, 1931, pp. 296–313.

6 "Friction Factors for Pipe Flow," by L. F. Moody, Trans. ASME, vol. 66, 1944, pp. 671–684.

7 "Friction in the Flow of Suspension," by E. G. Vogt and R. R. White, *Industrial and Engineering Chemistry*, vol. 40, 1948, pp. 1731–1738.

8 "Prediction of Pressure Drop During Forced-Circulation Boiling of Water," by R. C. Martinelli and D. B. Nelson, Trans. ASME, vol. 70, 1948, pp. 695–702.

9 "Pressure Drop Accompanying Two-Component Flow in a Closed Conduit with Various Liquids and Air," by E. G. Thomsen, MS thesis, University of California, Berkeley, Calif., 1941.

10 "Pressure Drop and Heat Transfer Accompanying Two-Component, Two-Phase Flow in Horizontal Pipes," by R. W. Ravenscroft, MS thesis, University of California, Berkeley, Calif., 1943.

Appendix 1

Momentum Pressure Change

The change of momentum of the air may be neglected. The change of pressure due to liquid momentum change over a finite length of tube can be expressed as

$$\Delta P_M = \frac{G_L}{g}(V_{L2} - V_{L1}) \quad [17]$$

where V_{L1} and V_{L2} are the liquid velocities at points a finite distance apart, and G_L is the liquid-mass velocity in lb/sec ft^2 of tube cross section. The velocity of the liquid is related to the mass velocity by the equation

$$V_L = G_L/R_L\rho_L \quad [18]$$

Substituting Equation [18] in Equation [17] gives

$$\Delta P_M = \frac{G_L^2}{g\rho_L}\left(\frac{1}{R_{L2}} - \frac{1}{R_{L1}}\right) \quad [19]$$

The gas velocity may be expressed as

$$V_G = G_G/(1 - R_L)\rho_G \quad [20]$$

Combining Equations [18] and [20] results in

$$\frac{V_L}{V_G} = \frac{G_L}{G_G}\frac{(1 - R_L)}{R_L}\frac{\rho_G}{\rho_L} \quad [21]$$

If, during a particular flow, V_L/V_G is assumed constant over the finite distance Δx and if the liquid is assumed incompressible, then

$$\left(\frac{1}{R_L} - 1\right) \propto \frac{1}{\rho_G}$$

Hence treating the air as a perfect gas

$$\left(\frac{1}{R_L} - 1\right) \propto \frac{1}{P}$$

where P is the absolute pressure. If P and $P - \Delta P_{TP}$ denote the pressures at points 1 and 2, respectively, then

$$\left(\frac{1}{R_{L1}} - 1\right)\Big/\left(\frac{1}{R_{L2}} - 1\right) = \frac{P - \Delta P_{TP}}{P} = 1 - \frac{\Delta P_{TP}}{P} \quad [22]$$

Hence

$$\frac{1}{R_{L2}} - \frac{1}{R_{L1}} = \left(\frac{1}{R_{L2}} - 1\right)\frac{\Delta P_{TP}}{P} \quad [23]$$

Substituting Equation [23] in Equation [19] gives

$$\Delta P_M = \frac{G_L^2}{g\rho_L}\left(\frac{1}{R_{L2}} - 1\right)\frac{\Delta P_{TP}}{P} \quad [24]$$

This equation is now applied to the smooth-tube tests with maximum flow rates. The properties are

$$G_L = 570 \text{ lb/sec ft}^2, \quad P = 3480 \text{ psf}, \quad \Delta P_{TP} = 86.5 \text{ lb/(sq ft) ft}$$

$$R_L = 0.351, \quad \rho_L = 62.3 \text{ pcf}$$

$$\Delta P_M = \frac{570^2}{32.2 \times 62.3}\left(\frac{1}{0.351} - 1\right)\frac{86.5}{3480} = 7.45 \text{ lb/(sq ft)ft}$$

i.e., the momentum pressure change is approximately 8.5 per cent of the friction pressure change.

Appendix 2

Pressure-Drop Relationships

The liquid pressure drop, calculated assuming the liquid flows alone in the tube, may be expressed by the equation

$$\Delta P_L = \frac{\lambda_S \Delta L V_{LP}^2 \rho_L}{2gD} \quad [25]$$

where the friction factor for smooth tubes (for example) during turbulent flow is of the form

$$\lambda_S = C'/N_{RLP}^n \quad [26]$$

The liquid Reynolds number is

$$N_{RLP} = \rho_L D V_{LP}/\mu_L \quad [27]$$

Substituting Equation [27] in Equation [26] we obtain

$$\lambda_S = C'(\mu_L/\rho_L D V_{LP})^n \quad [28]$$

If Reynolds number is defined with respect to the liquid velocity during two-phase flow, the friction factor becomes

$$\lambda' = C'(\mu_L/\rho_L D V_L)^n \quad [29]$$

The liquid velocity V_L during two-phase flow is related to the liquid velocity V_{LP} assuming the liquid flows alone in the tube, in the equation

$$A_L V_L = A_P V_{LP} \quad [30]$$

where A_L and A_P are the liquid and tube cross sections, respectively. Hence

$$V_L = V_{LP}\frac{A_P}{A_L} = V_{LP}/R_L \quad [31]$$

Combining Equations [28], [29], and [31]

$$\lambda_S = \lambda'/R_L^n \quad [32]$$

Substituting Equations [31] and [32] in Equation [25], we find

$$\Delta P_L = \frac{\lambda' \Delta L V_L^2 \rho_L}{2gD} R_L^{2-n} \quad [33]$$

As n for the smooth and the galvanized tubes are 0.25 and 0.125, respectively, combining Equation [33] with Equations [9] and [10] leads to the form

$$\Delta P_{TP} = 0.8\frac{\lambda' \Delta L V_L^2 \rho_L}{2gD} \quad [11]$$

Discussion

Ovid Baker.[5] The authors have developed a good approach to rough pipes for two-phase flow. Their methods of analysis have been applied to data on 7.75 and 10.136-in-ID pipe in annular and slug flow at high pressures[6] in smooth pipe. The $\lambda/\lambda s$ ratio for these tests was essentially 1.0. The data are plotted in Fig. 15 together with the authors' Equation [1] for smooth pipe.

The comparison shown in Table 2 for the authors' method of correlation and that of the discusser[6] indicates about equal deviation from experimental results if $\bar{X}$-values less than 0.4 are excluded. Most gas-distillate-field two-phase gathering lines do fall in the excluded range. The pressure drop in these economically important pipelines can be predicted by methods already published.

In the range of $\bar{X} = 0.2$ to 2.5, data for a 7.75-in-ID line in annular flow may be correlated by the equation

$$\frac{\Delta P_{TP}}{\Delta P_L} - 1 = \frac{37.27}{(\bar{X})^{1.759}}$$

For 10.136-in. pipe the value of $m = 1.759$ is satisfactory but the C value is slightly less than 37.27.

For higher $\bar{X}$-values in the 7.75-in-ID pipe, the flow pattern changes to slug flow and is correlated approximately by the equation

$$\frac{\Delta P_{TP}}{\Delta P_L} - 1 = \frac{7}{(\bar{X})^{0.394}}$$

In this case the C-value for 10.136-in. pipe would be considerably higher. This equation does not correlate the data as well be-

[5] Magnolia Petroleum Company, Dallas, Tex.

[6] "Designing for Simultaneous Flow of Oil and Gas," by Ovid Baker, *Oil and Gas Journal*, July 26, 1954, pp. 185–90, 192, 195.

Table 2

Experiment no.	Type flow pattern	ID of pipe, in.	Length of test section, ft	ΔP_L calc.	$\frac{\Delta P_{TP}}{\Delta P_L} - 1$	$\bar{X}$	ΔP_{TP} experimental, psi	ΔP_{TP} authors' equation	Per cent deviation from equation	ΔP_{TP} Baker[6]	Per cent deviation from equation
1	Annular	7.750	11317	0.0307	617.9	0.200	19	3.92	−79	22.32	+17
2	Annular	7.750	11317	2.00	15.0	1.73	32	27.40	−14	42.57	+33
3	Annular	7.750	11317	1.265	6.9	2.51	10	12.27	+23	13.07	+31
7	Annular	10.136	41333	0.0322	589.0	0.198	19	4.15	−78	14.26	−25
8	Annular	10.136	41333	2.08	12.9	1.69	29	28.91	0	21.58	−26
									+23		+27
					Average per cent deviation for annular flow				−14		−26
4	Slug	7.750	11317	2.85	4.96	3.77	17	19.24	+13	17.16	+1
5	Slug	7.750	11317	0.848	6.07	5.04	6	4.45	−26	5.97	−1
6	Slug	7.750	11317	1.92	4.21	7.45	10	7.49	−25	11.44	+14
9	Slug	10.136	41333	0.964	13.5	2.53	14	9.25	−34	18.41	+32
10	Slug	10.136	41333	2.878	7.34	3.79	24	19.28	−20	29.50	+23
11	Slug	10.136	41333	1.74	8.19	5.08	16	9.05	−43	21.03	+31
12	Slug	10.136	41333	1.937	8.29	7.91	18	7.17	−60	18.30	+2
									+13		+17
					Average per cent deviation for slug flow				−34		−1

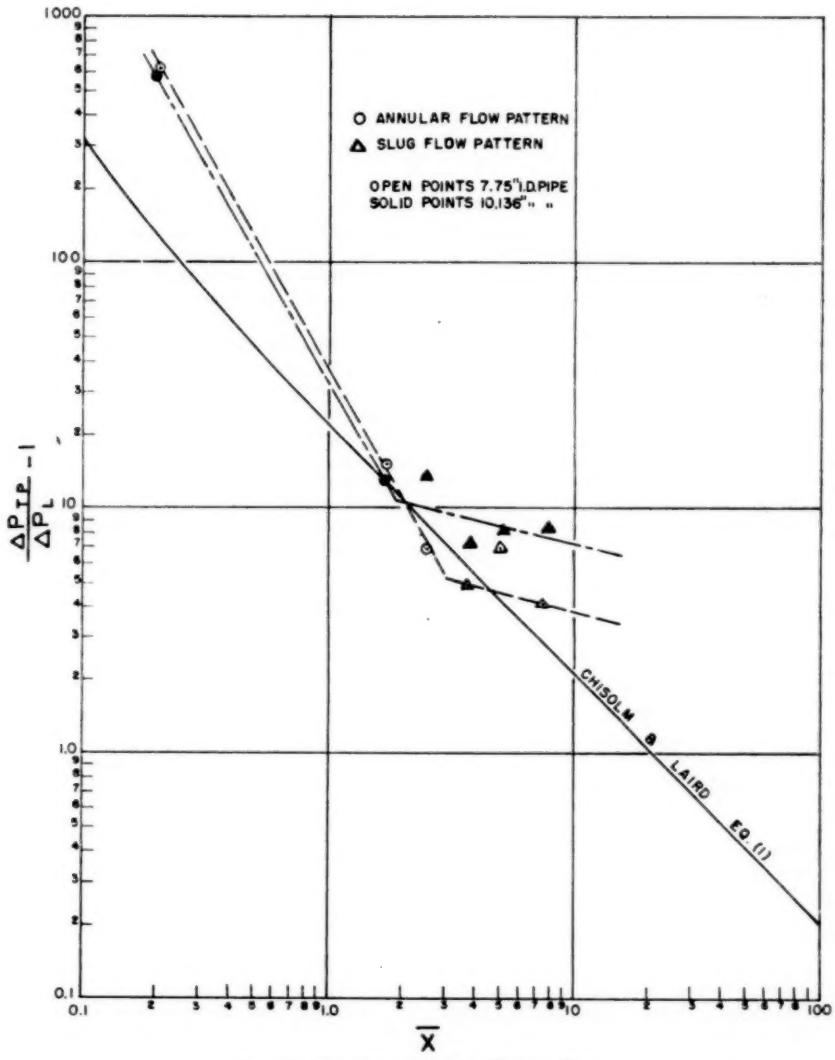


FIG. 15 TWO-PHASE FLOW IN SMOOTH PIPE

cause the liquid-mass velocity, which is a major variable in slug flow, is not included in the correlation.

The large differences in the values of C and the change in values of m when the flow pattern changes from annular to slug indicates that the flow pattern must be considered especially in industrial-size pipelines.

H. S. ISBIN.[7] The authors have treated a very complex problem and have been able to reduce their data in the form of some unusually good correlations.

At the University of Minnesota, we have been working on two-phase flow, but unlike the authors, our studies have been confined to a one-component, steam-water system. The purpose of this discussion is to stimulate interest in new directions, or at least, to leave the impression that the Martinelli-type correlations are not "universally applicable."

[7] Associate Professor, Department of Chemical Engineering, University of Minneapolis, Minneapolis, Minn.

The writer does indeed contest the statement ". . .that the Reynolds number, the tube roughness, and the flow rates uniquely determine the characteristics of isothermal two-phase flow." Perhaps it is understood that other parameters are included such as system pressure and flow geometry, incorporating specification and arrangement of flow channel as well as other fluid properties which might affect the distribution of the phases in the flow channel.

It has been the writer's experience that the Martinelli-type correlations cannot be used with confidence in the prediction of steam-water pressure drops. The work has been reported in a preliminary manner[8] and a paper is now being prepared on pressure drops for the adiabatic flow of steam-water mixtures covering the following range of variables:

[8] "Two-Phase Pressure Drops," by H. S. Isbin, R. H. Moen, and D. R. Mosher, Government Research Report, AECU-2994, Nov., 1954.

System pressure	25 to 1415 psia
Flow rates	454 to 4350 lb/hr
Qualities	0 to 100 per cent
Pipe diameters	0.484 and 1.062 in.

Although an empirical correlation of the data has been found the writer believes that his own work and the contributions in the authors' paper are only intermediate answers. Further measurement on over-all pressure-drop measurements will not be as fruitful as measurements on the phase distribution in the flow channels. To the writer's knowledge, no method of approach has yet been described in the literature which will yield the new insight which is so long overdue.

Several other comments are offered. The term "liquid saturation" has been used in the paper and corresponds to the use of "liquid fraction" or (1—void fraction). It is suggested that we standardize the usage to "liquid fraction." The momentum calculation in the Appendix is somewhat arbitrary in that specific assumptions have been made to confine the momentum-pressure change to an interpreted change in the liquid fraction. The authors should state whether, over the range of flow rates measured, the humidification of the air and the resulting temperature changes did not produce significant momentum-pressure drops.

Authors' Closure

Mr. Baker's additional data and calculations in Table 2 are most interesting. They seem to indicate that tube size is less important than flow pattern. The knees in the curves of Fig. 15 at transition between flow types are similar to those in many other publications. Much of the scatter of two-phase flow is caused by such transitions, but they were not noticeable in the present research. A good method of treating transition problems would be valuable, but remains to be found. The main purpose of this paper, however, was to consider the effects of pipe-wall roughness.

As Mr. Isbin suggested, the statement to which he took exception needs proper interpretation. The units of the flow rate were not specified, but were considered as properly reflecting the system pressure, which in combination with the Reynolds numbers must control the flow geometry. Martinelli-type correlations are frequently little better than guesses when applied to arbitrary systems for which they were not developed. The authors agree that present methods of correlation would warrant little confidence for water vapor-liquid systems over such wide ranges of variables. The authors subscribe to the use of the expression "liquid fraction" instead of "liquid saturation." The expression "volume fraction" instead of "saturation" is also recommended.

Laminar Flow Over an Enclosed Rotating Disk

By S. L. SOO,[1] PRINCETON, N. J.

Laminar flow over an enclosed rotating disk was studied to reduce the inconsistency between previous theoretical and experimental results. Unlike the case of a disk in an infinite fluid medium, the friction-moment coefficient of the enclosed disk is proportional to Re^{-1} in the laminar range and $Re^{-1/4}$ in the turbulent range. The latter is accurate for the range of disk diameter to gap ratios between 200 and 50. The former correlation, instead of being an approximation according to simple shear as has been suggested, is shown to be true even when recirculation exists. The deviation from Re^{-1} relation in the laminar range is shown to be due to the inertia effect of recirculation. The significance of inertia effects in addition to that due to centrifugal force has been pointed out and corrections have been presented. Radial outflow has been shown to be more effective than radial inflow for turbine-disk cooling. It has been shown that rotation of the shroud provides an added sealing effect for a centrifugal machine.

Nomenclature

The following nomenclature is used in the paper:

a_n, a_n' = coefficients in series expansion of W
A = Re_z^2
b_n, b_n' = coefficients in series expansion of V
c_n, c_n' = coefficients in series expansion of U
C = constant of integration
D = diameter of disk, ft
k_s = friction-moment coefficient
m = net flow rate, slug/sec
M = friction moment, ft-lb (one side)
n = rpm
p = pressure, psf
r = radial co-ordinate or radius as defined, ft
r_s = disk radius, ft
Re_z = Reynolds number based on $\omega z_0^2/\nu$
Re_r = Reynolds number based on $\omega r_s^2/\nu$ or nr_s^2/ν as stated
u = radial component of velocity, fps
U = dimensionless velocity as defined
v = peripheral component of velocity, fps
V = dimensionless velocity as defined
w = axial component of velocity, fps
W = dimensionless velocity as defined
z = axial co-ordinate, ft
z_0 = gap between disk and housing, ft
ζ = dimensionless axial co-ordinate
ζ_2 = $1 - \zeta$
θ = angle, rad
μ = viscosity, lb/ft sec
ν = kinematic viscosity, ft²/sec
ρ = density, slug/cu ft
τ = shear stress, lb/ft²
ψ = stream function as defined, ft/sec
ω = angular velocity, rad/sec
$'$, $''$, etc. = first, second, etc., derivatives of the variable with respect to ζ

Introduction

Studies of the problem of friction and heat transfer for a disk rotating in an infinite fluid medium already have been made thoroughly (1–6).[2] Experimental measurements of disk friction in a finite housing have shown that the case of an enclosed disk is quite different from the case of a disk in an infinite medium, although many of the theoretical correlations follow the trend of that for an infinite system (7–9). The study reported here was made to obtain a better understanding of the phenomena associated with the enclosed rotating disk.

Problems of importance in engineering which can be generalized in the case of flow over an enclosed rotating disk include windage losses, air cooling of turbine disks (10), pedestal bearings with center feed of lubricant (11), and leakage flow over the shroud of a centrifugal pump or compressor. The analysis presented in this paper is restricted to the case of incompressible laminar flow; an approximate solution for turbulent flow has been given in the discussion.

Only a limited amount of numerical computation was possible in connection with the present study. However, the method is suitable for solutions by automatic computers in cases where the applications justify their use.

Formulation of the Problem

The system shown in Fig. 1 consists of a disk or plate rotating at constant angular velocity ω in an incompressible fluid and situated at a distance z_0 from a stationary plate or boundary. Symmetrical with the axis of rotation the fluid flows radially inward or outward at a mass rate of flow m.

The equation of continuity and the momentum equations for steady flow may be expressed in cylindrical co-ordinates as

$$\frac{1}{r}\frac{\partial ru}{\partial r} + \frac{\partial w}{\partial z} = 0 \quad \ldots\ldots [1]$$

$$u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} - \frac{v^2}{r} = -\frac{1}{\rho}\frac{\partial p}{\partial r} + \nu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) - \frac{u}{r^2} + \frac{\partial^2 u}{\partial z^2}\right] \quad \ldots\ldots [2]$$

$$\frac{u}{r}\frac{\partial rv}{\partial r} + w\frac{\partial v}{\partial z} = \nu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v}{\partial r}\right) - \frac{v}{r^2} + \frac{\partial^2 v}{\partial z^2}\right] \quad \ldots\ldots [3]$$

[1] Associate Professor of Mechanical Engineering, Princeton University. Assoc. Mem. ASME.

Contributed by the Fluid Mechanics Committee of the Hydraulics Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 4, 1957. Paper No. 57—SA-28.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

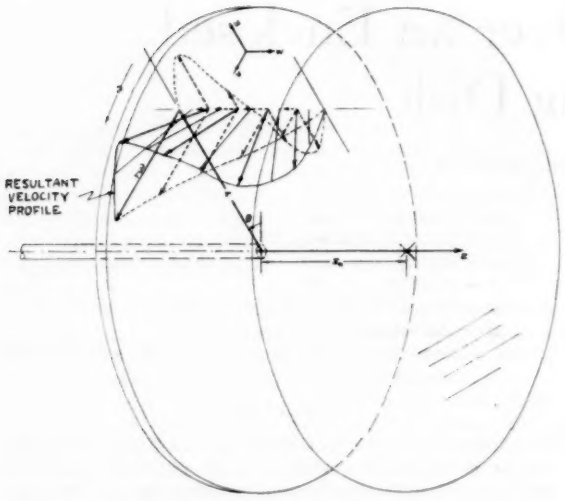


FIG. 1 CO-ORDINATE SYSTEM OF AN ENCLOSED ROTATING DISK, SHOWING VELOCITY PROFILE AT RADIUS r AT THE CONDITION OF NO NET RADIAL FLOW

$$u\frac{\partial w}{\partial r} + w\frac{\partial w}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial z} + \nu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial w}{\partial r}\right) + \frac{\partial^2 w}{\partial z^2}\right] \quad [4]$$

with the following boundary conditions

$$\left.\begin{aligned} &u(r, 0) = 0; \quad u(r, z_0) = 0; \quad 2\pi\rho\int_0^{z_0} ru\,dz = m; \\ &v(r, 0) = r\omega; \quad v(r, z_0) = 0; \\ &w(r, 0) = 0; \quad w(r, z_0) = 0; \end{aligned}\right\} \quad [5]$$

$m > 0$ for radial outward flow;
$m < 0$ for radial inward flow

The flow rate m may be assumed to be small in all the practical cases mentioned previously.

APPROXIMATE SOLUTION FOR SMALL m

Introducing the usual boundary-layer approximations, Equations [2] to [4] can be simplified to

$$u\frac{\partial u}{\partial r} + w\frac{\partial u}{\partial z} - \frac{v^2}{r} = -\frac{1}{\rho}\frac{\partial p}{\partial r} + \nu\frac{\partial^2 u}{\partial z^2} \quad [6]$$

$$\frac{u}{r}\frac{\partial rv}{\partial r} + w\frac{\partial v}{\partial z} = \nu\frac{\partial^2 v}{\partial z^2} \quad [7]$$

$$\frac{\partial p}{\partial z} = 0 \quad [8]$$

Equation [1] can be satisfied by the stream function ψ defined by

$$w = \frac{1}{r}\frac{\partial r\psi}{\partial r}, \quad u = -\frac{\partial \psi}{\partial z} \quad [9]$$

Further, in cases where m may be considered small, the substitutions

$$\psi = \frac{\omega z_0^2}{\nu} r\omega z_0 W(\zeta) + \frac{m}{2\pi\rho}\frac{U(\zeta)}{r} \quad [10]$$

and

$$v = r\omega V(\zeta) \quad [11]$$

where

$$\zeta = z/z_0 \quad [12]$$

into Equations [6] and [7] enables separation of the variables. This leads to

$$W^{iv} + 2AWW''' - 2VV' = 0 \quad [13]$$

$$V'' + 2AW'V - 2AWV' = 0 \quad [14]$$

$$U^{iv} - 2AWU''' - 2AWU'' = 0 \quad [15]$$

where

$$A = \left(\frac{\omega z_0^2}{\nu}\right)^2 = \mathrm{Re}_s^2 \quad [16]$$

and Re_s is the Reynolds number of the system based on z_0. The velocity components are given by

$$u = -\mathrm{Re}_s r\omega W' - \frac{mU'}{2\pi\rho z_0 r} \quad [17]$$

$$w = 2\mathrm{Re}_s \omega z_0 W \quad [18]$$

$$v = r\omega V \quad [19]$$

together with the boundary conditions

$$\left.\begin{aligned} &W(0) = W(1) = 0 \\ &V(0) = 1, \; V(1) = 0 \\ &U'(0) = U'(1) = 0 \\ &U(1) - U(0) = 1 \end{aligned}\right\} \quad [20]$$

FIRST-ORDER APPROXIMATION

A set of simplified solutions can be obtained when one takes Re_s as arbitrarily small. This is equivalent to neglecting all inertia effects except that due to centrifugal force as in reference (7). In other words, if one takes A equal to zero in Equations [13] to [15], the solution reduces to

$$U = 3\zeta^2 - 2\zeta^3 \quad [21]$$

$$U' = 6\zeta - 6\zeta^2 \quad [22]$$

$$V = 1 - \zeta \quad [23]$$

$$W = -\frac{\zeta^2}{20} + \frac{7}{60}\zeta^3 - \frac{\zeta^4}{12} + \frac{\zeta^5}{60} \quad [24]$$

$$W' = -\frac{\zeta}{10} + \frac{7}{20}\zeta^2 - \frac{\zeta^3}{3} + \frac{\zeta^4}{12} \quad [25]$$

The results are shown by solid lines in Figs. 2, 3, 4, and 5 for comparison with the solution given later for finite values of A.

Equation [23] shows that simple shear exists in the peripheral direction across the gap z_0. Essentially, this simplified solution constitutes a first-order approximation. The results, however, are good enough for most engineering requirements.

The resisting moment M due to friction can be obtained by integrating the shearing unit stress over the area of the disk. The shearing stress is

$$\tau = \mu\frac{\partial v}{\partial z} = -\mu\frac{r\omega}{z_0} \quad [26]$$

and the moment of one side is given by

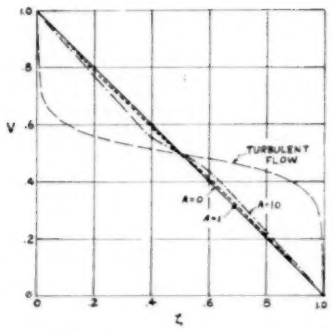


FIG. 2 DIMENSIONLESS PERIPHERAL VELOCITY V ($v = r\omega V$, $A = \mathrm{Re}_r^2$)

$$M = 2\pi \int_0^{r_0} \tau r^2 dr = \frac{\pi}{2} \frac{\mu\omega}{z_0} r_0^4 \quad \ldots\ldots\ldots [27]$$

The moment coefficient of friction (8) is

$$M \Big/ \left[\int_0^{r_0} \int_0^{2\pi} \frac{1}{2} \rho (rw)^2 r \, r \, d\theta \, dr \right]$$

$$k_r = \frac{M}{\pi\omega^2\rho r_0^5/5} = \frac{5}{2\mathrm{Re}_z} \left(\frac{z_0}{r_0}\right) = \frac{5}{4\mathrm{Re}_r} \left(\frac{D}{z_0}\right) \ldots [28]$$

where

$$\left.\begin{aligned} \mathrm{Re}_r &= \omega r_0^2/\nu \\ \text{and} \qquad D &= 2r_0 \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots [29]$$

Equation [28] has been plotted against $\mathrm{Re}_r = nr_0^2/\nu$ in Fig. 6 for comparison with the experimentally determined curve (7, 8). Other theoretical curves (2, 7) are included for comparison. Reference (7) neglected momentum terms other than that due to centrifugal force but considered the axial component to be of similar order as the peripheral component.

SECOND-ORDER APPROXIMATION

Including the effect of finite values of A in Equations [13], [14], and [15], while still considering only small values of m, may be considered as a second-order approximation. The solution of this case requires series expansions from $\zeta = 1$ to $\zeta = 1/2$. Equilibrium requires that

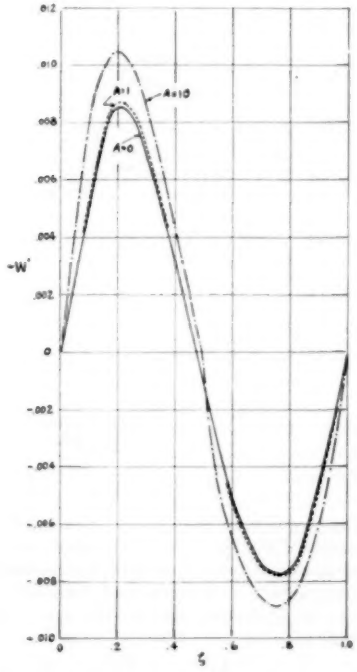


FIG. 4 DIMENSIONLESS RADIAL VELOCITY W'. RADIAL VELOCITY DUE TO ROTATION $= -\mathrm{Re}_z\, r\omega W'$, $A = \mathrm{Re}_r^2$

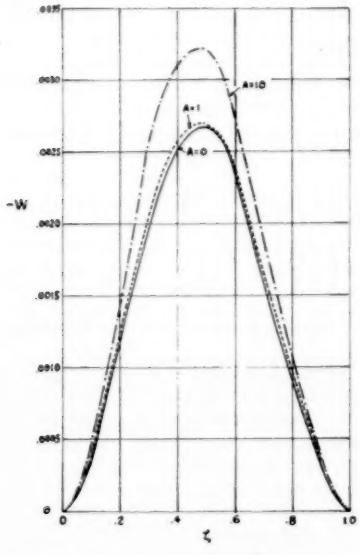
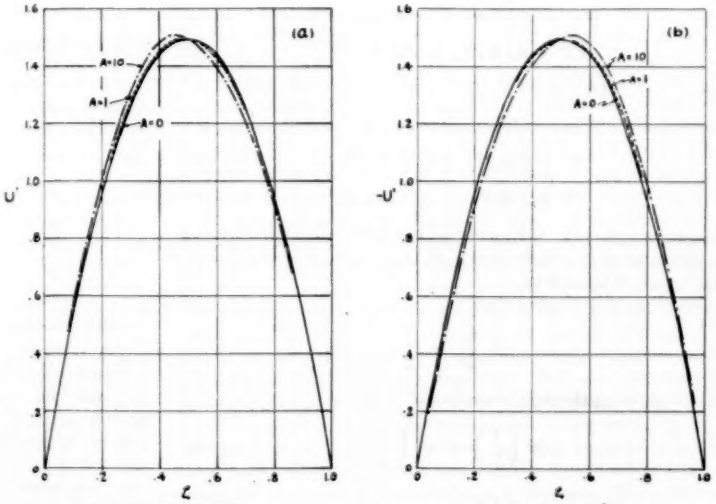


FIG. 3 DIMENSIONLESS AXIAL VELOCITY W ($w = 2\mathrm{Re}_z\, \omega\, z_0\, W$, $A = \mathrm{Re}_r^2$.)



(a) Radial outflow (b) Radial inflow

FIG. 5 DIMENSIONLESS RADIAL VELOCITY U'. RADIAL VELOCITY DUE TO NET FLOW $= -\dfrac{mU'}{2\pi\rho z_0 r}$, $A = \mathrm{Re}_r^2$

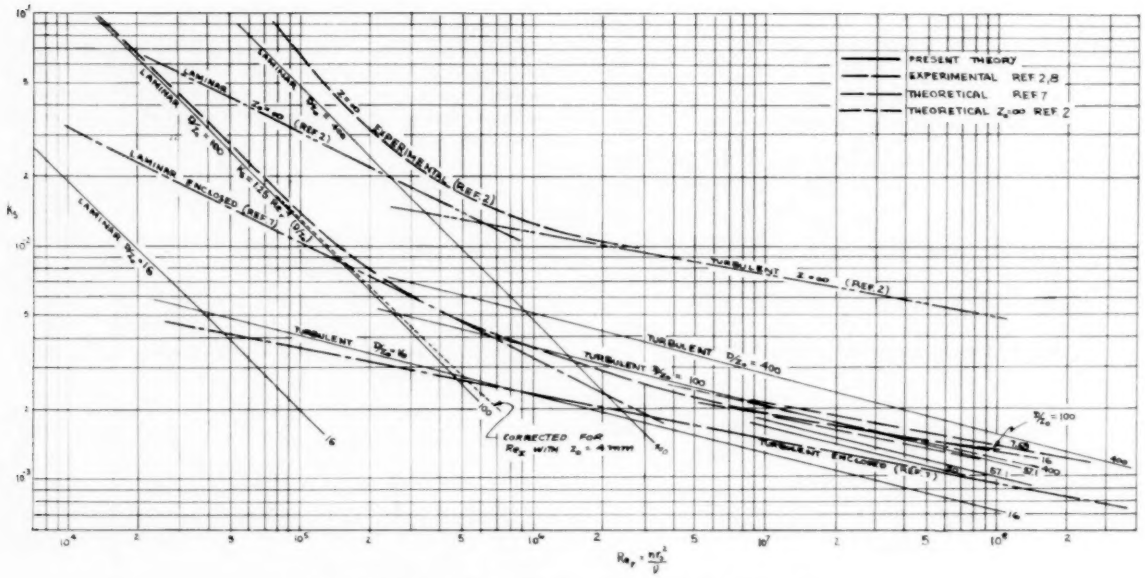


FIG. 6 COMPARISON OF THEORETICAL AND EXPERIMENTAL RESULTS
[n = rpm, k_3 is defined as the ratio of friction moment on one side and $(\pi\omega^2\rho r_3^5/5)$.]

$$V'(0) = V'(1) \quad [30]$$

since the same torque must be transmitted even if a disk and its housing are reduced to narrow rings of control volume of radius dr. In other words, with small mass flow, the angular momentum flux from the space between the disks must be small and therefore the friction torques on the two disks must be equal. This suggests the introduction of

$$\zeta_2 = 1 - \zeta \quad [31]$$

for $1/2 < \zeta < 1$, into Equations [13], [14], and [15] which become

$$W_2^{iv} + 2AW_2W_2''' + 2V_2V_2' = 0 \quad [32]$$

$$V_2'' - 2AW_2'V_2 + 2W_2V_2' = 0 \quad [33]$$

and

$$U_2^{iv} + 2AW_2U_2''' + 2AW_2'U_2'' = 0 \quad [34]$$

The subscript 2 denotes functions of ζ_2. Using series expansions

$$W = \Sigma\, a_n\zeta^n, \quad W_2 = \Sigma\, a_n'\zeta_2^n \quad [35]$$

$$V = \Sigma\, b_n\zeta^n, \quad V_2 = \Sigma\, b_n'\zeta_2^n \quad [36]$$

Equations [13], [14], [32], and [33] can be solved simultaneously by determining the coefficients a_n, b_n, a_n', and b_n' from the following boundary conditions:

Equation [30] requires that

$$\Sigma\, \frac{b_n}{2^n} = \frac{1}{2} \quad [37]$$

At $\zeta = \zeta_2 = 1/2$ the requirements are

$$W\left(\frac{1}{2}\right) = W_2\left(\frac{1}{2}\right) \quad [38]$$

$$W'\left(\frac{1}{2}\right) = W_2'\left(\frac{1}{2}\right) \quad [39]$$

$$W''\left(\frac{1}{2}\right) = W_2''\left(\frac{1}{2}\right) \quad [40]$$

$$W'''\left(\frac{1}{2}\right) = W_2'''\left(\frac{1}{2}\right) \quad [41]$$

Equations [15] and [34] can be expressed as

$$U''' - 2AWU'' = C \quad [42]$$

$$U_2''' + 2AW_2U_2'' = C \quad [43]$$

where C is the constant of integration which must be determined.

From the known functions W and W_2, Equations [15] and [34] can be solved for U and U_2 by using the series expansion

$$U = \Sigma\, c_n\zeta^n, \quad U_2 = \Sigma\, c_n'\zeta_2^n \quad [44]$$

with the boundary conditions

$$U(0) = 0, \quad U_2(1) = 1 \quad [45]$$

$$U\left(\frac{1}{2}\right) = U_2\left(\frac{1}{2}\right) \quad [46]$$

$$U'\left(\frac{1}{2}\right) = U_2'\left(\frac{1}{2}\right) \quad [47]$$

$$U''\left(\frac{1}{2}\right) = U_2''\left(\frac{1}{2}\right) \quad [48]$$

The coefficients of the foregoing series expansions are presented in the Appendix.

Numerical solutions of the simultaneous Equations [37] to [41] and [46] to [48] determine the coefficients of the series expansion of U, V, and W. Determination up to b_4 calls for the solution of five simultaneous equations of second order. As the value of A increases, more terms have to be taken into account which increases the order of the algebraic equations accordingly.

Figs. 2 to 5 show the results of solutions for the cases $A = 1$ (dotted line) and $A = 10$ (dash-dot line). The peripheral component of velocity in this case is represented by

$$v = r\omega(1 + b_1\zeta + \ldots)$$

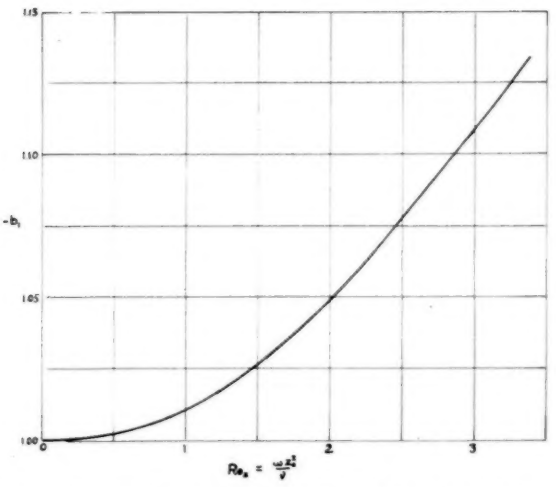


FIG. 7 COEFFICIENT b_1 OR INERTIA CORRECTION FACTOR OF MOMENT COEFFICIENT (SHEAR AT THE WALL = $b_1 r\ \omega/z_0$)

Values of the coefficient b_1 are shown in Fig. 7. These are the correction coefficients of k_s for various values of Re_s within the laminar-flow range; at large values of A, this correction for moment coefficient becomes increasingly significant.

DISCUSSION

1 The boundary-layer solution obtained in the foregoing appears justified by comparison with the experimental results, Fig. 6, for small spacing. The first-order approximation for $Re_s \rightarrow 0$ differs from the solution in reference (7) only in that the latter accounts for $\partial^2 w/\partial z^2$. Equations [17] to [19] show $\partial^2 w/\partial z^2$ to be small when compared to $\partial^2 u/\partial z^2$ and $\partial^2 v/\partial z^2$ since w is of the order of the gap z_0 while u and v are of orders of the radius r. The emphasis of $\partial^2 w/\partial z^2$ while disregarding the momentum contribution other than that due to centrifugal force leads to the analogous condition of a disk in an infinite medium. The flow condition over an enclosed disk, however, is basically different from the case of an infinite medium; any effort in trying to get similarity between the two cases is not likely to lead to significant result. For wider disk spacing the correction of inertia effect becomes increasingly significant, the limiting case is when the spacing is infinite, in which case an exact solution is given in reference 4, and the inertia effects are accounted for in an exact manner.

2 In the case of the first-order approximation, the fluid is under simple shear v across the gap. The result obtained here is a generalization of the simple case presented in reference (1). The centrifugal force induces the recirculation u outward at the disk surface and inward at the surface of the housing. The induced w being negative means flow from near the housing toward the disk.

3 The second-order approximations for finite values of Re_s show the significance of the inertia of the fluid in modifying the conditions of simple shear. The inertia effects tend to increase the shear at the wall, Fig. 2. The curve labeled turbulent flow in Fig. 2 shows a boundary layer based on the $1/7$ power velocity-distribution law. This indicates the manner by which inertia effects aid the transition from a laminar to a turbulent boundary layer through instability (5). The presence of a point of inflection has been suggested as a necessary and sufficient condition for instability (13). The effect of finite values of Re_s also can be seen in Fig. 4. In the limit, the increased effects of inertia and instability transform the motion u into that of turbulence with a thin boundary layer (7).

4 Simple peripheral shear across the gap is a very good approximation, Fig. 6, when the gap z_0 is very small compared to the radius r. The moment coefficient for this case is

$$k_s = \frac{5}{4\mathrm{Re}_r}\left(\frac{D}{z_0}\right)$$

In the experimental case mentioned previously, $Re_s = 0.418$ for water at $Re_r = \omega r_s^2/\nu = 10^6$, the correction for the second-order effect of Re_s is very small (dotted line in Fig. 6). Therefore Equation [28] is a very good approximation for the condition in pedestal bearings. Fig. 7 shows that the first-order approximation would lead to 15 per cent error when Re_s is equal to 3.8. Therefore a correction may be necessary in the case of gas-turbine disks and centrifugal-compressor shrouds.

On the basis of the $1/7$ power velocity-distribution law (2), it can be shown with the present method that, where the turbulence dissipation in the core is small when compared to that due to shear at the walls, the moment coefficient in the turbulent regime is given by, for turbulent boundary layer of half of the thickness of the spacing

$$k_s = 0.0206\ \mathrm{Re}_r^{-1}\left(\frac{D}{z_0}\right)^{1/6} \quad \ldots\ldots\ldots\ldots\ [49]$$

(Appendix) which also checks closely with the experimental results, Fig. 6. The independence of the gap in the transition from the laminar to the turbulent range as represented in reference (7) is due mainly to the similarly prescribed motion for all sizes of gap by neglecting momentum in the boundary layer other than that due to centrifugal force while considering viscous forces. It is inconceivable that, for similar disk, housing, fluid, and speed, similar motion occurs whether the gap is 0.01 in. or 1 in. Comparison with experimental results of reference (8) shows that reference (7) provides reasonable over-all approximations only for $D/z_0 < 50$; i.e., for wide spacing. On the other hand, Equation [49] tends to give too high a value of k_s at large D/z_0 (>200) and too low at small D/z_0 (<50). Hence reference (7) provides a good approximation for wide gap, while for small gap, other velocity laws should be taken in the derivation of Equation [49] because, below certain values of D/z_0, the solution should reduce to the case of $z_0 = \infty$ (1, 4).

A proper explanation of the continuity of the experimental curve is probably that for the range $Re_r > 10^6$ represented in Fig. 6, transition from laminar to turbulent boundary layer exists in the space between the disk and the housing. In other words, laminar motion always exists for a solid disk in a housing. The similarity of the dimensionless correlation of reference (7) to that of the case of a disk in an infinite medium (1) can be attributed physically as due to overemphasis of the axial motion while neglecting momentum other than that due to centrifugal force. The axial motion is one of the main motions in the latter case but in the former case, its effect is small when the gap is small. The proportionality to $Re^{-1/2}$ in the laminar case of reference (7) should not be attributed to recirculation; recirculation occurs even for very low Reynolds numbers. The approximation in Equation [49], however, is good only when turbulent dissipation in the core is small when compared to that at the walls.

From similarity between friction and heat transfer (12), the heat transfer from an enclosed rotating disk should follow a trend similar to the friction characteristics. However, the measurements reported in reference (14) do not substantiate this fact.

5 The accuracy of both of the foregoing approximations can be seen by substituting Equations [21] to [25] into Equations [2] and [3]. The criterion is that the quantity

$$\frac{30m}{\pi\rho z_0 r^2 \omega \mathrm{Re}_s} < 1$$

for high accuracy. Fortunately this is true in almost all practical cases.

6 When the net flow is small, the recirculation-velocity component is but little affected by the net flow. W and U are independent in the case $\mathrm{Re}_s \to 0$ so the actual value of u can be obtained by superposition. For instance, in the case of radial outflow, there is no recirculation for any radius smaller than r_1 given by

$$r_1^2 = \frac{45m}{\pi\omega\rho z_0 \mathrm{Re}_s} \quad \ldots\ldots\ldots\ldots\ldots [50]$$

The radial flow velocity at $r = r_1$ is

$$u = \frac{1}{60} r_1 \omega \mathrm{Re}_s \left[\frac{2U'}{3} - 60\,W'\right] \quad \ldots\ldots\ldots [51]$$

The trend is shown in Fig. 8. Similarly, for radial inflow there is no recirculation below r_1'

$$(r_1')^2 = \frac{30m}{\pi\rho z_0 \omega \mathrm{Re}_s}$$

where m is the strength of radial inflow. The possible cases of flow over a rotating disk are shown diagrammatically in Fig. 9 when there is no casing at the outer edge. It is seen that, without net flow, recirculation takes place over the whole radial dimension. With net flow less than the foregoing limiting values, local recirculation takes place in the form of a torus.

7 For substantial values of Re_s, U' is modified by W as shown in Fig. 5. The distortion will be more pronounced at higher values of A. In the case of radial outflow, such modification tends to increase the radial shear rate at the disk side while, in the case of radial inflow, such modification tends to increase the radial shear rate at the housing side.

8 According to both items 7 and 6, in general, radial outflow over the disk increases the shear at the surface of the disk and decreases the shear at the housing in comparison with the case of no net radial flow. This trend is quite favorable in the source flow cooling of a gas-turbine disk. The cooling of a radial-inflow turbine disk by radial inflow of cool air will be less effective. This effect can be seen further in Fig. 9.

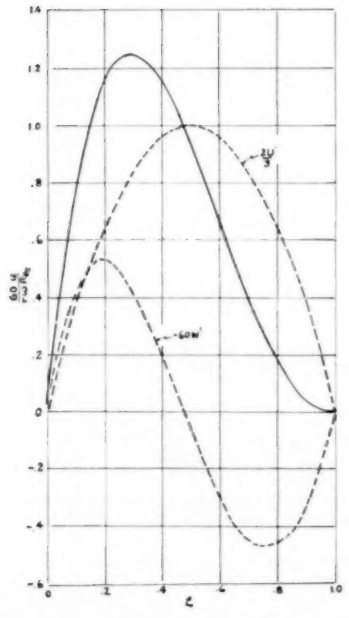


FIG. 8 LIMITING CASE OF SOURCE FLOW AND ROTATION AT RADIUS r_1

9 The average radial pressure variation between any radius r_0 and r can be shown to be

$$\frac{p - p_0}{\frac{1}{2}\rho r^2\omega^2} \leqq -\frac{3}{\pi}\left(\frac{m}{\rho z_0 \omega r^2}\right)\left(\frac{1 - 0.078968\,\mathrm{Re}_s^2}{\mathrm{Re}_s}\right)\ln\left(\frac{r}{r_0}\right) + \frac{3}{10\pi^2}\left(\frac{m}{\rho z_0 \omega r^2}\right)^2\left(\frac{r^2}{r_0^2} - 1\right) + \frac{2}{5}(1 + 0.001095\,\mathrm{Re}_s^2)\left(1 - \frac{r_0^2}{r^2}\right) \quad \ldots\ldots [52]$$

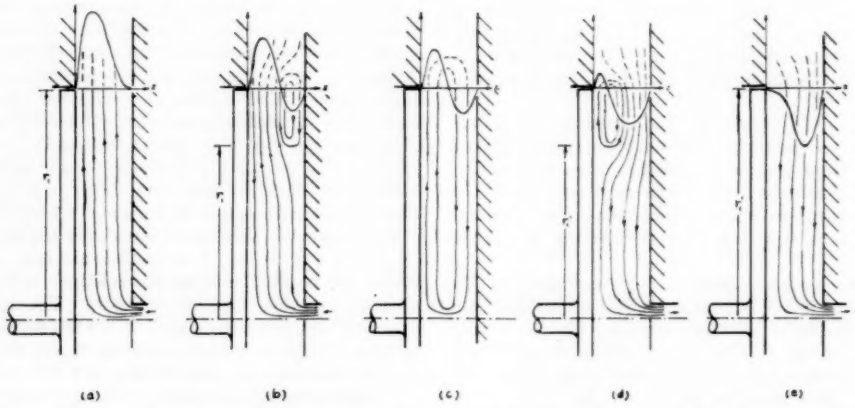


FIG. 9 RADIAL FLOW OVER ROTATING DISK AT VARIOUS FLOW CONDITIONS WHEN THERE IS NO CASING AT THE OUTER EDGE

[(a) Radial outflow, with limiting flow at rim; (b) radial outflow, flow quantity less than limiting value at rim; (c) no net flow, pure recirculation; (d) radial inflow, flow quantity less than limiting value at rim; (e) radial inflow, with limiting flow at rim.]

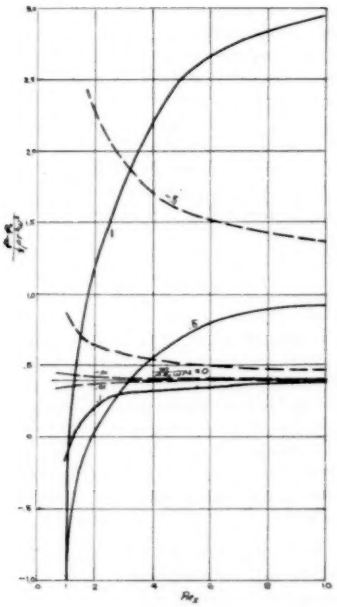


FIG. 10 AVERAGE PRESSURE DIFFERENCE BETWEEN DISK AND HOUSING FOR $r/r_0 = 10$
($m > 0$ for radial outflow, $m < 0$ for radial inflow.)

The first term on the right-hand side accounts for friction imparted on radial net flow, the second term accounts for diffusion (radial outward flow) or acceleration (radial inward flow), and the third term represents the radial pressure variation due to centrifugal effect. For very small Re_s

$$\frac{p - p_0}{\frac{1}{2}\rho r^2\omega^2} = -\frac{3}{\pi \mathrm{Re}_s}\left(\frac{m}{\rho z_0 \omega r^2}\right)\ln\frac{r}{r_0} + \frac{3}{10\pi^2}\left(\frac{m}{\rho z_0 \omega r^2}\right)^2\left(\frac{r^2}{r_0^2} - 1\right) + \frac{2}{5}\left(1 - \frac{r_0^2}{r^2}\right) \quad [53]$$

The last term of Equation [53] represents the minimum average radial pressure distribution over an enclosed rotating disk. Equation [53] was plotted as shown in Fig. 10, for both inward ($m < 0$) and outward flow ($m > 0$). It can be seen that rotation always tends to reduce the leakage flow of a centrifugal pump or compressor, but the minimum inlet pressure required to carry out radial outflow cooling really depends on the combination of the parameters Re_s and $m/\rho z_0 \omega r^2$. In the case of radial outflow, as in cooling a gas-turbine disk, the diffusion in the radial flow tends to raise the static pressure, while fluid friction tends to decrease the pressure due to the increase in velocity (15). The data shown in Fig. 10 for large values of $m/\rho z_0 \omega r^2$ tends to be overoptimistic due to instability of flow under the adverse pressure gradient. The centrifugal force tends to attenuate the boundary layer at the disk surface. At the housing, the recirculation tends to aggravate boundary-layer separation.

CONCLUSIONS

1 The solution obtained here for an enclosed rotating disk is valid. The friction moment coefficients

$$k_s = \frac{5}{4\mathrm{Re}_s}\left(\frac{D}{z_0}\right) \text{ for laminar flow}$$

$$k_s = 0.0206\ \mathrm{Re}_s^{-1/4}\left(\frac{D}{z_0}\right)^{1/4} \text{ for turbulent flow}$$

are quite accurate for small gaps between the disk and housing.

2 For large values of Re_s which are usually associated with wide gaps, the friction moment obtained from simple shear theory can be 15 per cent too small at values of Re_s of the order 4. The second-order effect of Re_s must be considered if a closer approximation is required.

3 The method developed for including the second-order effect of Re_s can be carried out with automatic computers.

4 Radial outflow enables very effective cooling of a disk surface.

5 Rotation provides an added sealing effect of the shroud of a centrifugal machine.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Dr. Robert C. Dean, Jr., Ingersoll-Rand Corporation, Phillipsburg, N. J., and reviewers for their suggestions.

BIBLIOGRAPHY

1 "Grenzschicht Theorie," by H. Schlichting, Verlag and Druck G. Braun, Karlsruhe, Germany, 1951.

2 "Modern Developments in Fluid Dynamics," by S. Goldstein, Oxford University Press, London, England, vol. 2, section 164, 1938, p. 367.

3 "Forced Flow Against a Rotating Disc," by D. M. Hannah, Reports and Memoranda, N. 2772 (British), 1952.

4 "Heat Transfer by Laminar Flow From a Rotating Plate," by K. Millsaps and K. Pohlhausen, *Journal of the Aeronautical Sciences*, vol. 19, February, 1952, p. 127.

5 "Theoretical Study of Turbulent Transition of a Rotating Circular Disc," by I. Shibuya, *Reports of the Institute of High Speed Mechanics*, Tōhōku University, Japan, vol. 1, 1951, p. 27.

6 "Heat Transfer From a Rotating Plate," by R. L. Young, ASME Paper No. 54—SA-51.

7 "Der Reibungswiderstand rotierender Scheiben in Gehausen," by F. Schults-Grunow, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 15, July, 1935, pp. 191–204.

8 "Versuche über Scheibenreibung," by K. Pantell, *Forschung auf dem Gebiete des Ingenieurwesens*, vol. 16, no. 4, 1949, pp. 97–108.

9 "The Influence of Viscosity on Centrifugal Pump Performance," by A. T. Ippen, ASME Paper No. 45—A-57.

10 "The Determination of Temperature Distribution in Gas Turbine Rotor Bodies and Cylinders by the Electrolytic Tank Method," by H. Baumann, *The Brown-Boveri Review*, vol. 40, May-June, 1953.

11 "Development and Preliminary Tests of a Rotating Viscosimeter," by H. W. Emmons and S. L. Soo, Harvard University, Cambridge, Mass., May, 1952.

12 "Heat Transfer," by M. Jakob, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1949, p 438.

13 "On the Stability of Three-Dimensional Boundary Layers with Application to the Flow Due to a Rotary Disk," by N. Gregory, J. T. Stuart, and W. S. Walker, Symposium on Boundary Layer Effects in Aerodynamics, NPL England, 1955, HMSO.

14 "Experimental Cooling of Radial Flow Turbines," by E. N. Petrick and R. D. Smith, ASME Paper No. 54—A-245.

15 "Theory of Laminar Flows in Convergent or Divergent Pipes," by H. Ito, *Reports of the Institute of High Speed Mechanics*, Tōhōku University, Japan, vol. 3, December, 1950.

Appendix

DETERMINATION OF COEFFICIENTS OF SERIES EXPANSION

The coefficients determined from Equations [13] and [15] are

$a_0 = 0,$	$b_0 = 1$
$a_1 = 0,$	b_1
$a_2,$	$b_2 = 0$
$a_3,$	$b_3 = -2Aa_2/3$
$a_4 = b_1/12,$	$b_4 = -Aa_3(3 + b_1)/6$

$a_5 = b_1^2/60$, $b_5 = -Ab_1(1 + 6a_2b_1)/30$

$a_6 = -Aa_2(1 + 3a_2)/90$, $b_6 = -A(3b_1^2 + 8Aa_2^2)/180$

$a_7 = -Aa_2(3 + 9a_3 + 8b_1)/630$, $b_7 = -A^2a_2(-1 - 18a_2 + 30a_3 - 5a_2b_1)/315$

The coefficients determined from Equations [32] and [33] are

$a_0' = 0$, $b_0' = 0$

$a_1' = 0$, $b_1' = -b_1$

a_2', $b_2' = 0$

a_3', $b_3' = -b_3$,

$a_4' = 0$, $b_4' = -b_4$

$a_5' = -b_1'^2/60$ $b_5' = -b_5$

$a_6' = Aa_2^2/30$, $b_6' = -b_6$

$a_7' = (3Aa_2'a_3' + 2b_1'b_2')/210$, $b_7' = -b_7$

The basic unknowns b_1, a_3, a_3, and a_2' and a_3' can be determined by simultaneous solution of Equations [37] to [41]. Equations [45] to [48] provide the information needed to determine the unknown coefficients C, c_2, and c_2' in the following

$c_0 = 0$, $c_0' = 1$

$c_1 = 0$, $c_1' = 0$

c_2, c_2'

$c_3 = C/6$, $c_3' = -C/6$

$c_4 = 0$, $c_4' = 0$

$c_5 = Ac_2a_3/15$, $c_5' = -Ac_2'a_2'/15$

$c_6 = A(2c_2a_3 + Ca_2)/60$, $c_6' = -A(2c_2'a_3' + Ca_2')/60$

$c_7 = A(c_2b_1 + 6Ca_3)/630$, $c_7' = -A(c_2'b_1' + 6Ca_3')/630$

Derivation of Moment Coefficient in Turbulent Range

Assuming the validity of $^1/_7$ velocity law (1) at the wall and taking the velocity of the core as one half of the disk velocity, the shear stress at the wall is given by

$$\tau_0 = 0.03955\ \rho \left(\frac{r\omega}{2}\right)^{7/4} \nu^{1/4} z_0^{-1/4}$$

The moment coefficient

$$k_s = \frac{2\pi \int_0^{r_s} \tau r^2 dr}{\pi \rho \omega^2 r_s^5/5} = 0.0206 \left(\frac{D}{z_0}\right)^{1/4} \mathrm{Re}_r^{-1/4}$$

Discussion

R. E. Nece[3] and J. W. Daily.[4] The writers wish to compliment the author on his contribution in generalizing the solution for the case of simple laminar flow between a rotating smooth disk and a closely spaced fixed boundary. His findings and conclusions regarding the details of velocity distribution, flow circulation, and heat-transfer effects appear physically reasonable and it is expected they can be verified by suitable experiments.

The author also discusses turbulent flow and the regime he describes as the transition to turbulence. The writers believe this needs further clarification, and, in view of their own experiments with smooth and rough enclosed disks which are currently under way at the M.I.T. Hydrodynamics Laboratory, are prompted to make the following comments.

In this paper, the author's discussions for both laminar and turbulent flow are concerned only with the case of close axial clearance between a rotating disk and a stationary boundary. In such instances, the flow regime is characterized by converged boundary layers; that is, the boundary-layer thickness on both rotating and stationary walls is constant over the radius and equal to $z_0/2$. For both laminar and turbulent flow, there is a second regime obtained with decreasing D/z_0 (increasing axial clearance) such that for constant

$$\mathrm{Re}_r = \frac{\omega r_s^2}{\nu}$$

the boundary layers, which are a function of Re_r and therefore have the same thickness for different z_0 values, will be relatively thinner leaving a "core" of finite width which rotates at approximately one half the velocity of the disk. Furthermore, for a turbulent boundary layer, the thickness will vary over the radii of both the disk and stationary boundary.

In general, it should be expected that as Reynolds number, Re_r is increased from a low value, the successive flow regimes encountered may be:

1 Laminar flow with converged boundary layers.
2 Laminar flow with rotating core.

Then, following the transition from laminar to turbulent flow with an attendant initial increase in boundary layer thickness:

3 Turbulent flow with converged boundary layers.
4 Turbulent flow with rotating core.

Of course, the transition from the laminar to turbulent regime is not likely to be sharp because, since the local Reynolds number varies as radius squared, turbulence no doubt occurs first at the periphery and works progressively inwards. This has been demonstrated experimentally for a free disk rotating in air in the author's reference (13).

The author's Equation [29] applies to regime No. 1, his Equation [49] to No. 3. The relations from reference (7) of the author's Bibliography plotted on the author's Fig. 6 represent solutions for regimes Nos. 2 and 4. These relations in the author's notation are

2nd regime $$k_s = \frac{1.06}{(\mathrm{Re}_r)^{1/2}} \quad [54]$$

4th regime $$k_s = \frac{0.0248}{(\mathrm{Re}_r)^{1/4}} \quad [55]$$

with

$$\mathrm{Re}_r = \frac{\omega r_s^2}{\nu}$$

It will be noted that Equations [54] and [55] do not predict an effect of axial clearance.

[3] Assistant Professor of Hydraulics, Massachusetts Institute of Technology, Cambridge, Mass.

[4] Professor of Hydraulics, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

* It should be noted that the author's Fig. 6 is plotted using

$$\mathrm{Re}_r = \frac{\mathrm{rpm} \times r_s^2}{\nu}$$

whereas ω instead of rpm is used in the Reynolds numbers in his equations. Parenthetically, it might be noted that the equations corresponding to [54] and [55] for a "free disk" ($z_0 = \infty$), which in the author's notation are

$$k_s = \frac{1.53}{(\mathrm{Re}_r)^{1/2}} \text{ and } k_s = \frac{0.058}{(\mathrm{Re}_r)^{1/5}}$$

respectively (using ω in the Reynolds number), appear to be plotted incorrectly in the author's Fig. 6.

The question arises as to the magnitude and range of Reynolds numbers over which these several regimes might exist, and also, of course, as to the accuracy of the equations for calculating the disk-friction torque. Briefly, the facts of the situation seem to be the following:

1 For turbulent flow, all published torque formulas have been derived using the physical model of a rotating core, corresponding, therefore, to the high Reynolds-number range and regime No. 4.

2 Between the laminar flow, regime No. 2 and the high Reynolds-number turbulent-flow range, a "transition" curve is sometimes observed, although up to now regime No. 3 has not been explicitly tagged.

3 Contrary to the predictions of the previously published formulas, torque measurements by the writers for regimes No. 2 and No. 4, as well as published measurements for regime No. 4 (see author's Fig. 6 and reference 8) have indicated variations with the axial-clearance ratio such that as D/z_0 increases, the torque coefficient decreases.

4 The author's Equation [49] indicates increasing torque coefficient with increasing D/z_0.

In view of these observations, it appears that the range over which regime No. 3 may exist will depend on the axial clearance ratio, decreasing with decreasing values of D/z_0 (increasing clearance). M.I.T. experiments indicate for D/z_0 less than about 30, this regime may not be distinguished from a laminar-to-turbulent transition range of the usual sort. As to the lack of agreement between published formulas and experiment for the wide clearance regimes Nos. 2 and 4, an important shortcoming of all the theoretical analyses has been the omission of the effect of wall friction at the cylindrical portion of the enclosing housing on the loss of momentum and kinetic energy. It is concluded that an adequate theoretical development for torque equations must be based on a more realistic physical model than has been used heretofore.

R. C. Dean, Jr.[5] The author is to be complimented on presenting an analysis of the laminar disk-friction and boundary-layer problem which, within the knowledge of the writer, is more complete and authentic than any other in the literature for the case of a small gap between the disk and casing. The solution yields flow patterns which show evidence of circulation in the clearance gap which bears some resemblance to the Taylor-Gortler rings found in the gap between a rotating cylinder and stationary coaxial cylindrical casing. The work of Kaye and Elgar[6] shows that these vortexes have a significant influence on heat transfer, and presumably also on friction, in the cylindrical case and would be expected similarly to influence the heat transfer from and friction of a rotating disk in a closely spaced casing by promoting a vigorous mixing of the flow.

Kaye and Elgar's investigation demonstrated that the flow between cylindrical walls could occur in four regimes depending upon the mass flow and rotative speed of the inner cylinder. These regimes are characterized by the flow being laminar or turbulent with or without the vortexes. The writer suggests that the problem of the flow between a disk and housing might be investigated profitably by using visualization and hot-wire techniques to reveal the detailed space-time characteristics of the flow. These characteristics may aid immeasurably, as in Elgar's case, to explain the variations in integrated parameters as shown in the author's Fig. 6.

M. A. Santalo[7] and W. A. Wilson.[8] There are two distinct approaches to the solution of fluid-flow problems in engineering devices. This paper is a creditable example of an elegant mathematical treatment based on the general form of the Navier-Stokes equations. These fundamental relationships have been simplified by "introducing the usual boundary-layer approximations," and further simplified by restricting attention to the case of very small Reynolds numbers and very small integrated radial flows. The method then yields precise solutions for the cases consistent with these assumptions.

The second approach is also classic; it involves positing the essential kinematic characteristics of the solution, frequently a one-dimensional description or symmetry in velocity profiles. The kinematic parameters of the hypothesis are then adjusted to satisfy certain gross limitations embodied in the continuity, momentum, and energy relationships. Solutions thus obtained are susceptible to indefinite refinement by applying the method simultaneously to smaller and smaller subdivisions of the process. In the limit, of course, the latter method coincides with the former, but in practice and in spirit the two methods are quite distinct. The second is based on the assumption of the possibility of a direct physical apprehension of the phenomenon and the first on confidence in being able to cull from the general differential equations those terms of negligible consequence.

The first method holds out the promise of definitive solutions, either analytical or numerical. In fluid-mechanics problems we usually must be content with severe restrictions on the range of applicability of these solutions. In the present instance the addition to our understanding of flow over a rotating disk is limited to the correction factor presented in Fig. 7. The author suggests the applicability of this correction to flows over gas-turbine disks and centrifugal-compressor shrouds. As a matter of fact, the Re_r range, 0 to 3.5, covered by Fig. 7 probably covers the laminar-flow regime for which the curve is valid, but typical applications in turbines and compressors lie far outside it. (For example if $\omega(D/2) = 1000$ fps, $z_0 = 1/16$ in. and the fluid is air at 200 F, then $Re_r = 78.5$.)

Of greater practical importance is the turbulent regime. The author presents some conclusions about this regime in his Discussion. Although we are not led in detail through the reasoning by which he comes to these conclusions it is fairly clear that this reasoning is of the second kind; i.e., he posits the $^1/_7$ power velocity-distribution "law." Such use of analogy can be very useful indeed. A paper by Hisao Jimbo[9] examines several aspects of the problem of flow over a rotating disk starting with friction laws developed by Schultz-Grunow in reference (7) of the paper, and the present paper. His conclusions which were reasonably substantiated by experiment have several practical implications.

There are two specific questions which the writers would like to direct to the author:

1 Are the slight asymmetries of the velocity profiles in Figs. 3, 4, and 5 attributable to some physical phenomenon or to the simplifications introduced into the basic equations? In particular, for the case $A = 0$, $m = 0$, and $D/z \gg 1$ the asymmetry indicated is very surprising.

2 What is the definition of "average pressure difference" as used in the caption of Fig. 10? If the assumption $\partial p/\partial z = 0$ ap-

[5] Head, Advanced, Engineering Department, Ingersoll-Rand Company, Phillipsburg, N. J. Mem. ASME.

[6] "Modes of Adiabatic and Diabatic Fluid Flow in an Annulus with an Inner Rotating Cylinder," by Joseph Kaye and E. C. Elgar, Research Laboratory of Heat Transfer in Electronics, Massachusetts Institute of Technology, Cambridge, Mass., Report No. RLHTE-13, February 15, 1957.

[7] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

[8] Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

[9] "Investigation of the Interaction of Windage and Leakage Phenomena in a Centrifugal Compressor," ASME Paper No. 56—A-47.

plies then $p = f(r)$ only and there would seem to be nothing to average. For the case of $m = 0$ and the symmetrical profiles of Fig. 2 one can conclude that at $z = z_0/2$

$$\frac{p - p_0}{\frac{1}{2}\rho r^2\omega^2} = 0.25$$

not 0.4 as indicated in Fig. 10.

A final comment is a school-masterish plea for more consideration of the readers in the preparation of curves. Fig. 6 which presumably relates the theory presented to other theories and to experimental work is virtually unintelligible to the writers.

Author's Closure

The author wishes to thank the discussers for their interest and criticism.

The four flow regimes pointed out by Professors Nece and Daily are certainly enlightening. While all these conditions tend to occur to a given disk, depending on the physical dimensions and properties, these are the four predominating cases. Their suggestion of the study including wall friction at the cylindrical portion of the enclosing housing is a worthy one in an effort toward a better understanding of disk friction problem. No doubt their calculations on the cases of free disks were right as a direct conversion, but the author has made a further conversion, based on the definition of moment coefficient given by Equation [28] in order to make comparison with the experimental results presented in reference (8).

The remarks contributed by Dr. Dean are valuable and instructive in supplementing this present study. His suggestion for experimental work will certainly lead to significant understanding.

The restriction as suggested by Professors Santalo and Wilson is really not there; as long as the flow is laminar, the method presented is not restricted to $\mathrm{Re}_s = 3.5$, nor $\mathrm{Re}_s = 78.5$. Even if calculations were carried out to the latter figure, one can always point out to an example in which Re_s is a larger quantity. In fact, the last paragraph of Introduction suggested these possibilities. The following might serve to answer their specific questions:

1 The viscosity profiles in Fig. 3 are not symmetric because centrifugal force always acts radially outward. The axial components in Fig. 4 are also not symmetric; Figs. 3 and 4 are related by continuity. The case of $A = 0$ in Fig. 5 is symmetric because it is simply Poiseuille flow.

2 The simplification as presented by Equation [8] means variation of pressure in the axial direction is negligible as compared to variations in the radial direction. In other words

$$\frac{\partial p}{\partial r} \simeq \frac{d\bar{p}}{dr}$$

where $\bar{p}$ means averaging in the axial direction, or Equation [52] was obtained from Equation [6] in the following manner

$$-\frac{1}{\rho}\frac{d\bar{p}}{dr} = \frac{\partial}{\partial r}\left[\frac{1}{z_0}\int_0^{z_0} u^2 dz\right] + \frac{1}{z_0}\int_0^{z_0} w\frac{\partial u}{\partial z}\,dz - \frac{1}{r} \times \left[\frac{1}{z_0}\int_0^{z_0} v^2 dz\right]$$

In their calculation in arriving at $(p - p_0)\Big/\left(\frac{1}{2}\rho r^2\omega^2\right) = 0.25$, they have neglected contributions from the radial component of velocity.

With regard to their final comment, the author wishes to assure the readers that better pictures will be prepared next time.

Influence of Various Grinding Conditions Upon Residual Stresses in Titanium

By P. A. CLORITE[1] and E. C. REED,[2] EAST HARTFORD, CONN.

Grinding conditions may be expected to affect residual surface stresses from any particular grinding operation. Stresses were measured in titanium test bars, surface-ground with various wheels, speeds, grinding fluids, down-feeds, and crossfeeds. Results suggest that, with suitable precautions, titanium alloys may be ground under either "near-normal" or "low-speed" conditions with acceptable grinding ratios and with low residual stresses.

Introduction

OVER the past few years several investigators have been concerned with various phases of residual stresses in steel, but little work has been done on residual-stress distributions in titanium resulting from various conditions of surface grinding. Tarasov (1)[3] is one of the few who has done grinding work in this field. This paper presents results of an initial study to determine the effects of surface-grinding conditions upon residual stresses in one titanium alloy. Additional work is contemplated in which correlation of residual stresses and fatigue life in titanium will be studied. This paper should not be construed to mean that the authors' company permits indiscriminate grinding of titanium or gives unqualified endorsement to titanium-grinding practices.

This work was conducted in two phases: (*a*) Grinding tests to determine what conditions lead to good grinding ratios, including investigations of various wheels, speeds, grinding fluids, and feeds; and (*b*) residual-stress tests in which stresses due to efficient grinding ratios were studied. All tests were made upon one lot of titanium-alloy specimens.

Experimental Procedure

Material and Equipment. 6 Al–4 V titanium rolled bars, with the following percentage composition, were used for these tests: 0.02 C, 0.27 Fe, 5.90 Al, 0.015 N_2, 3.88 V, 0.004 H_2, and remainder Ti. Hardness of the alloy was RC 30–32.

Grinding equipment included a reciprocating-table surface grinder, an 18-in. autocollimator (transit type), various 60-grit vitrified-bond grinding wheels (10 × ½ × 3), and various grinding fluids.

An optical interferometer (2) was used to determine residual stresses.

Grinding-Ratio Tests. Grind-test specimens were milled and finished by grinding to 3.0 × 6.0 × 2.0 in., with grain parallel to the 6.0-in. dimension, and vacuum annealed at 1250 F for 36 hr. After annealing, the specimens were fastened in a vise with a 3.0 × 6.0-in. face up. The vise was clamped to the table of the surface grinder. Before each test, the wheel was dressed clean and four 0.001-in. downfeed passes were made into the work to compensate for uncontrolled variables which may have resulted from dressing the wheel. The wheel diameter was measured with a micrometer, and one of two pieces of 0.062-in-thick plexiglas, held vertically in a fixture, was plunge-ground with the depth of grind being recorded. The test piece was then ground using the grinding conditions of the particular test. After down-feeding a total of 0.016 in. on the test piece, the spindle was raised to its initial height and the second piece of plexiglas was plunge-ground, using the depth of grind setting that was used on the first plexiglas piece. The difference in depth of cut in the two pieces of plexiglas was indicative of the amount of wheel "breakdown." This difference was measured with the autocollimator (a type of transit).

The thicknesses of all titanium test pieces were measured with a micrometer before and after each grinding test to determine the amount of stock removed. The grinding ratio, *G*-ratio (cubic inches of material removed per cubic inch of wheel wear), was computed. See Fig. 1 for typical values.

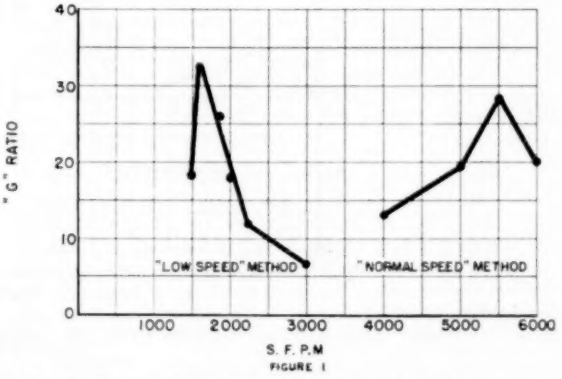


Fig. 1 Typical *G*-Ratios for Grinding 6 Al–4 V Titanium

In the course of testing, several grinding variables were studied, including type of wheel, wheel speed, grinding fluid, and feed. Tests were made by changing one variable at a time so that the effect of each could be evaluated separately.

Residual-Stress Tests. Interferometer strip specimens 3.0 × 0.75 × 0.3 in. were ground and vacuum-annealed at 1250 F for 36 hr. One 3.0 × 0.75-in. face of each specimen was polished by standard metallographic techniques to specular reflectance sufficient to produce an interference pattern when placed in contact with an optical flat and illuminated with mercury light. The opposite face of each specimen was surface-ground longitudinally in the surface grinder.

After grinding, each specimen was placed in the optical interferometer to obtain interference pattern caused by curvature of the polished face, resulting from removal of each increment of ground surface. Material was removed by etching with an aque-

---

[1] Methods Development Laboratory Engineer, Pratt & Whitney Aircraft Division, United Aircraft Corporation.

[2] Materials Development Laboratory Engineer, Pratt & Whitney Aircraft Division, United Aircraft Corporation.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 25, 1956. Paper No. 56—A-44.

ous solution of 30 per cent HNO_3 (sp gr 1.41) and 3 per cent HF (sp gr 1.24) by volume. Etching was done at room temperature with the entire specimen masked off except the ground surface. Increments removed were measured by determining weight loss from etching.

Uniaxial residual stresses were calculated from changes in longitudinal curvature of specimens as layers of ground surface were removed. This method has been used by Mattson (3) and Leaf (4) as well as by Reed (5).

Results

In evaluating grinding variables, tests for maximum *G*-ratios indicated that wheel type, grinding fluids, and speeds are not interchangeable between "low" and "normal-speed" methods, Fig. 1. The low-speed method used a table speed of 30 fpm, a monocrystalline (1 crystal/grain) aluminum-oxide wheel, aqueous potassium-nitrite grinding fluid, and wheel speed of 1800 sfpm for maximum *G*-ratio. The normal-speed method used a table speed of 30 fpm, a black-silicon-carbide wheel, sulphur-chlorinated, fatty-type-oil grinding fluid, and a wheel speed of 5500 sfpm for maximum *G*-ratio.

In the residual stress work, specimens also were surface-ground by a near-normal-speed method. This method used a table speed of 30 fpm, a black-silicon-carbide wheel, sulphur-chlorinated, fatty-type-oil grinding fluid, and a wheel speed of 4000 sfpm for slightly reduced maximum *G*-ratio.

In all residual-stress tests, grinding stresses were less than 0.005 in. deep, with peak tension stresses, in most cases, less than 0.001 in. below surface. Peak stresses ranged from 109,000 psi tension to 11,000 psi compression (Table 1).

Speed Effects. Low-speed grinding (1800 sfpm) resulted in lower maximum stresses than normal-speed grinding (5500 sfpm). Lowering wheel speeds of a particular method, other factors being constant, reduced residual stresses, Figs. 2 and 3.

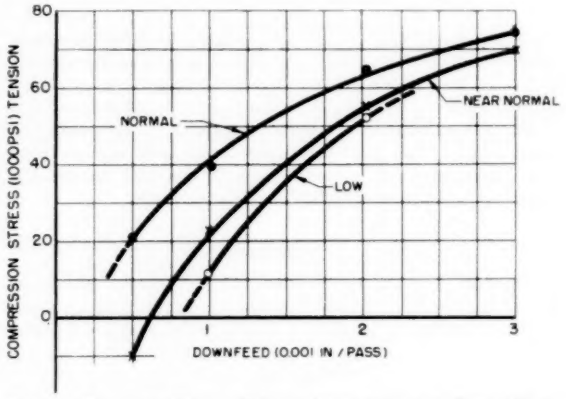


Fig. 2 Maximum Stress Versus Downfeed for Low, Near-Normal, and Normal-Speed Methods

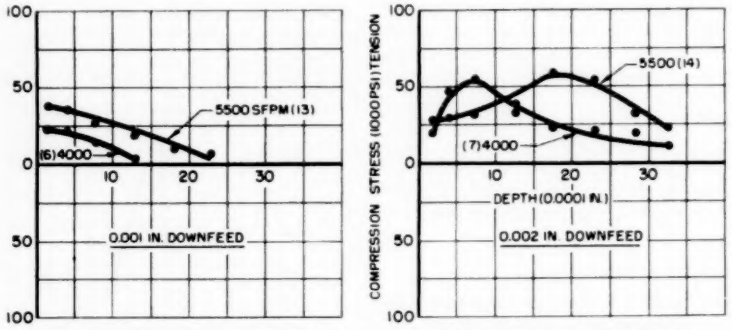


Fig. 3 Stress Distribution From 0.001 In. and 0.002 In. Downfeeds With Various Wheel Speeds

TABLE 1 TYPICAL RESIDUAL-STRESS DATA FROM SURFACE GRINDING OF TITANIUM (6 Al–4 V)

Test	Wheel	Wheel speed, sfpm	Grinding fluid	Downfeed, in./pass	Crossfeed, in./pass	Maximum stress, psi	Depth of stress, in.
1	Al_2O_3–K[a]	1800	KNO_2	0.001	0.050	11000	0.0012
2	Al_2O_3–K	1800	KNO_2	0.002	0.050	52000	0.0032
3	Black SiC–I	4000	Oil[b]	0.001	0.025	23000	0.0014
4	Black SiC–I	4000	Oil	0.002	0.025	57000	0.0028
5	Black SiC–I	4000	Oil	0.0005	0.050	—11000	0.0012
6	Black SiC–I	4000	Oil	0.001	0.050	22000	0.0014
7	Black SiC–I	4000	Oil	0.002	0.050	55000	0.0040
8	Black SiC–I	4000	Oil	0.003	0.050	70000	0.0050
9	Black SiC–I	4000	Oil	4 pass-0.003 1 -0.0005	0.050	22000	0.0050
10	Black SiC–I	5500	Oil	0.001	0.025	41000	0.0028
11	Black SiC–I	5500	Oil	0.002	0.025	70000	0.0048
12	Black SiC–I	5500	Oil	0.0005	0.050	21000	0.0014
13	Black SiC–I	5500	Oil	0.001	0.050	40000	0.0025
14	Black SiC–I	5500	Oil	0.002	0.050	64000	0.0044
15	Black SiC–I	5500	Oil	0.003	0.050	75000	0.0050
16	Black SiC–I	5500	Oil	3 pass-0.003 1 -0.001 1 -0.0005	0.050	78000	0.0040
17	Black SiC–I	5500	Oil	4 pass-0.003 1 -0.0005	0.050	62000	0.0050
18	Black SiC–I	5500	Oil	4 pass-0.003 2 -0.0005	0.050	56000	0.0040
19	Black SiC–K	5500	Oil	0.001	0.050	57000	0.0040
20	Black SiC–K	5500	Oil	0.002	0.050	109000	0.0040

[a] Al_2O_3 was monocrystalline, letter after dash is hardness.
[b] Oil was sulphur-chlorinated, fatty type.

TABLE 2 CROSSFEED EFFECTS ON RESIDUAL STRESSES FROM NORMAL AND NEAR-NORMAL-SPEED GRINDING

Wheel speed, sfpm	Crossfeed, in./pass	Max. stress, psi — Downfeed, in./pass 0.001	Max. stress, psi — Downfeed, in./pass 0.002	Depth of stress, in. — Downfeed, in./pass 0.001	Depth of stress, in. — Downfeed, in./pass 0.002
4000	0.025	23000 (3)	57000 (4)	0.0014	0.0028
4000	0.050	22000 (6)	55000 (7)	0.0014	0.0040
5500	0.025	41000 (10)	70000 (11)	0.0028	0.0048
5500	0.050	40000 (13)	64000 (14)	0.0025	0.0044

NOTE: Numbers in ( ) refer to stress test number.

Downfeed Effects. As grinding downfeeds decreased, other factors being constant, peak residual stresses decreased. Heavy downfeed passes followed by light downfeed passes reduced peak residual stresses from values obtained by heavy downfeed passes alone, Figs. 4 and 5.

Wheel Effects. Peak residual stresses were of the same magnitude from grinding with (i) aluminum-oxide wheels of the same hardness based on comparing (*a*) semi-friable (partially porous), (*b*) mixtures of regular (fused) and friable (porous), and (*c*) monocrystalline-type wheels, and (ii) black-silicon-carbide wheels of the same hardness, with other factors being constant. In any particular method, peak stresses were reduced as a softer wheel was used, Fig. 6.

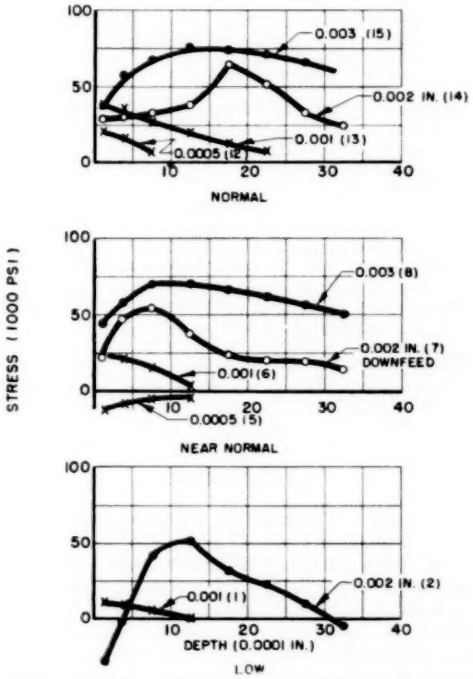


FIG. 4 STRESS DISTRIBUTION FROM LOW, NEAR-NORMAL, AND NORMAL-SPEED METHODS WITH VARIOUS DOWNFEEDS

TABLE 3 GRINDING FLUID EFFECTS ON *G*-RATIOS FROM NORMAL AND LOW-SPEED GRINDING[a]

Fluid base	Speed	*G*-ratio	Grinding fluid
Water	Low	25.0	Potassium nitrite
Water	Low	16.5	Sodium nitrite
Water	Low	2.8	High detergent, low lubricity
Water	Low	2.3	Poly-alkylene, glycol
Water	Low	1.5	Water emulsion
Oil	Normal	33.5	Sulphur-chlorinated, fatty
Oil	Normal	9.0	Chlorinated fatty
Oil	Normal	7.5	Sulphurized fatty, noncorrosive
Oil	Normal	3.5	Straight sulphurized, corrosive
Oil	Normal	3.0	Straight mineral

[a] 0.001 in./pass downfeed and 0.025 in./pass crossfeed.

Crossfeed Effects. Varying crossfeed from 0.025 in. per pass to 0.050 in. per pass had little effect on maximum stresses (Table 2).

Grinding-Fluid Effects. Maximum residual stresses were similar for low-speed grinding using the two potassium-nitrite grinding fluids tested. There were only slight differences in stresses for normal-speed grinding from use of the four sulphur-chlorinated, fatty-type oil grinding fluids tested. The sulphur-chlorinated, fatty-type oil used for residual-stress tests at normal speed had the most effective *G*-ratio characteristics. Among the water-base grinding fluids, potassium nitrites were the most effective. Table 3 lists the *G*-ratios of various oil and water-base grinding fluids tested in the grinding-ratio study.

DISCUSSION

Residual-stress results in this paper are based on uniaxial stresses. While the authors realize that grinding stresses are at least biaxial, a uniaxial method of determining stresses was used to reduce the amount of work in this first-phase study. It was felt that grinding conditions resulting in high stresses could be eliminated and that future testing could evaluate methods that have some possibility of producing an acceptable ground surface.

The grinding tests with their study of *G*-ratios revealed the fact

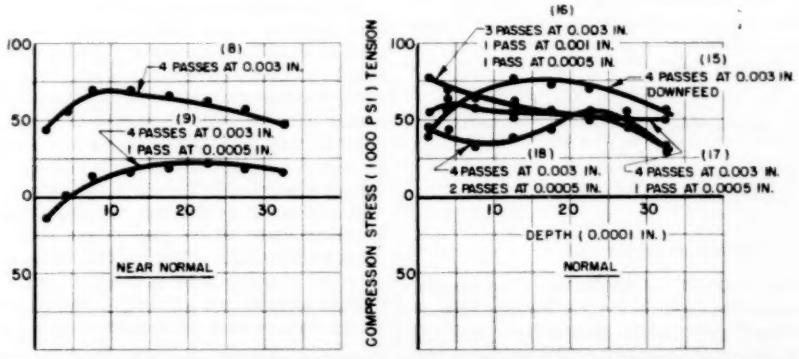


FIG. 5 STRESS DISTRIBUTION FROM NEAR-NORMAL AND NORMAL-SPEED METHODS WITH SUCCESSIVE PASSES OF VARIOUS DOWNFEEDS

that there are two definite possibilities of grinding titanium successfully—by the low-speed method and by the near-normal-speed method. Residual stresses from either are low. However, if the near-normal-speed method is used, alteration of grinding machines to obtain low-wheel speed is unnecessary.

It is interesting to note that Adenstedt (6) recommends the low-speed method for grinding titanium. These grinding tests also indicate that oils as well as nitrites can be used as grinding fluids in the low-speed method, but the authors favor nitrites as grinding machines must be shielded if oils are used.

Both *G*-ratio and stress results indicate that a sulphur-chlorinated, fatty-type oil grinding fluid can be used in near-normal-speed grinding. However, some grinding departments are reluctant to use oil because of a fire hazard from the spark stream. We believe that this fire hazard can be reduced by flooding the work with grinding fluid and keeping the grinding machine clean. The most dangerous element connected with this type of grinding is grinding sludge. Oil-soaked chips should not be left to accumulate or dry out, but should be disposed of regularly. The best rule to minimize fire is good housekeeping. It is believed by some fire underwriters that titanium chips offer less of a fire potential than do magnesium chips, as a low-velocity stream of water can be used to extinguish a titanium fire if one occurs.

In connection with grinding wheels, while no tests were made by the authors, there is a possibility that there may be some differences in any hardness of grinding wheel, from different manufacturers, as the mixtures of aluminum oxide or silicon carbide may vary or be in different form from wheel to wheel. For example a "J" wheel may be similar to a "K" wheel.

Hardness of the test titanium alloy was RC 30-32, but tests were not made to determine a possible change in surface hardness which might result from grinding. It is hoped that this can be accomplished in the future in connection with a fatigue program.

In the residual-stress-distribution data, stresses changed rapidly, in some cases from compression to tension. This was probably caused by cold-working and thermal effects. These rapid changes will need further study to determine their effect on fatigue life.

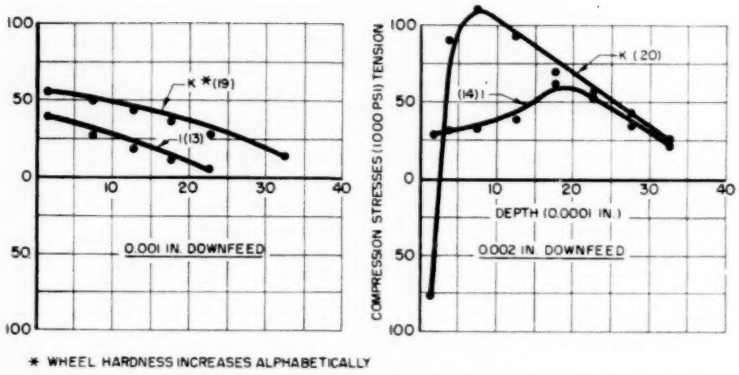


FIG. 6 STRESS DISTRIBUTION FROM 0.001 AND 0.002-IN. DOWNFEEDS USING NORMAL-SPEED METHODS WITH VARIOUS WHEEL HARDNESSES

CONCLUSIONS

Based upon the tests conducted, it is concluded that residual grinding stresses in titanium alloy can be held to low values and that *G*-ratios can be held to high values if any of the three following procedures is used:

Procedure	1	2	3
Wheel "I—K"	Aluminum oxide	Silicon carbide	Silicon carbide
Wheel speed, sfpm..	1800 "low"	4000 "near normal"	5500 "normal"
Grinding fluid.....	KNO_2	S-Cl fatty oil	S-Cl fatty oil
Downfeed, in./pass..	0.001	0.001	0.0005
Crossfeed, in./pass..	0.025–0.050	0.025–0.050	0.025–0.050
Table speed, fpm....	30	30	30

ACKNOWLEDGMENT

Thanks are extended to the Wheel Application, Metallurgical and Chemical Processing Groups of the Pratt & Whitney Aircraft Production Engineering Department for their co-operation, and Messrs. A. A. Taylor and O. P. Lowrey for their assistance and advice.

BIBLIOGRAPHY

1 "How to Grind Titanium," by L. P. Tarasov, *American Machinist*, vol. 96, Nov. 10, 1952, pp. 135–146.

2 "Application of Optical Interferometer to the Study of Residual Surface Stresses," by H. R. Letner, Proceedings, SESA, vol. X, no. 2, 1953, pp. 23–36.

3 "Method of Calculating the Residual Stress in a Simple Beam of Rectangular Cross-Section From Measurements of Its Longitudinal Curvature as Layers of the Material Are Removed," by R. L. Mattson, General Motors Corp., private communication, March, 1945.

4 "Techniques in Residual Stress Analysis," by W. Leaf, Proceedings, SESA, vol. IX, no. 2, 1952, pp. 133–140.

5 "Report on Residual Stresses—Aluminum Alloy Shot Peened Group 6 Specimens, SAE, ISTC Division IV," by E. C. Reed, MDL Report 1829, Pratt & Whitney Aircraft, May, 1955.

6 "Handbook on Titanium," by H. K. Adenstedt, AVCO, Technical Report 54-305, Part II, Wright Air Development Center, Sept., 1955, pp. V-2-25.

Discussion

L. P. TARASOV.[4] This paper is a most welcome addition to the literature on grinding stresses, all of which has thus far dealt solely with steels. As yet, nothing of a comparable nature has been published on machining stresses, and it is to be hoped that such information will become available for various materials and operations. To the extent that grinding conditions may be selected on the basis of residual-stress data, it is equally important to be able to do the same for machining conditions; otherwise, an unnecessarily high standard may be set up for grinding as compared to machining.

Only two tests in which aluminum-oxide wheels were used are listed in Table 1, and both were run with a nitrite solution. In view of the high compressive grinding stresses that Letner has found for hard steel ground with aluminum-oxide wheels and straight grinding oils, it would be most interesting to have some similar data for titanium ground at a low wheel speed. It is possible that appreciable compressive stresses could then be generated at higher rates of stock removal than with the silicon-carbide wheel used in Test 5.

It also would be desirable to obtain grinding-stress data for aluminum-oxide wheels used in the normal speed range, i.e., around 6000 sfpm. Although the resultant grinding ratios are very low, titanium is being precision ground under these conditions to some extent, primarily where the wheel is large relative to the work so that the diametral wheel wear is not excessive.

[4] Metallurgical Engineer, Research and Development Department, Norton Company, Worcester, Mass. Mem. ASME.

Granted that some titanium parts can be produced economically under these conditions, it would be well to know whether dangerously high tensile stresses are introduced in such operations.

In connection with the nitrite solutions that were used, were they made up from the pure compounds or were they commercially available nitrite-amine solutions? In either case, any information that can be presented about the concentrations used, either actual or in terms of the commercial products, would be helpful to those engaged in similar studies.

Authors' Closure

We wish to thank Dr. Tarasov for his remarks. We agree completely that there are many phases yet to be explained in regard to grinding and machining stresses in various metals. Our paper is only an initial step that we hope will lead to further studies.

In regard to the question about nitrite solutions, we used commercially available nitrite-amine solutions in a concentration of 1 part nitrite to 10 parts water.

A Tool-Work-Thermocouple Compensating Circuit

By K. J. TRIGGER,[1] R. K. CAMPBELL,[2] and B. T. CHAO[3]

A compensating circuit which facilitates the use of the tool-work thermocouple for interface-temperature measurements is presented. The *IR* drop due to the flow of thermoelectric current in a closed circuit is used to nullify the parasitic emf introduced by dissimilar lead materials attached to the cutting insert. Conditions necessary to achieve complete compensation are explained and test results indicating the reliability of the method are given.

Introduction

THE tool-work-thermocouple technique has been in use for over three decades in the study and measurement of metal-cutting temperatures. The interface between the tool and the work consists of a parallel network of junctions or bridges at all points of real contact. Each bridge constitutes a separate thermocouple hot junction. It has been shown[4,5] that the temperature distribution along the path of contact of the chip and tool is nonuniform and therefore the emf generated at the various junctions likewise will be nonuniform. The cooler junctions to some extent will act as "cold junctions" for all of the hotter bridges. Under such conditions the tool-work thermocouple will indicate a true arithmetic average only if the electrical resistances at all points of contact are equal.[5] Otherwise, the indicated emf may be greater or less than the arithmetic average depending upon the distribution of such resistances in relation to the temperature distribution. In the absence of quantitative information to the contrary it has been assumed that the junction resistances are approximately equal and that the tool-work thermocouple indicates the arithmetic average of the emf's (temperature) at all junctions. In the case of a "sharp" tool, i.e., flank wear not in excess of ~0.002 in., the interface is that at the chip contact. Under such conditions the term tool-chip thermocouple would appear to be more appropriate. In the presence of appreciable flank wear an additional interface at the flank-work contact is introduced and the indicated emf is affected by the additional junctions and their temperature distribution.

The conventional tool-work thermocouple as the term is used herein refers to a circuit in which the tool leg is a single continuous conductor such as a high-speed steel bar or to a composite of carbide tool materials, all components of which are identical in composition. In connection with the latter, it has been pointed out[6] that the calibration bar, the carbide contact rod, and the cutting insert must be pressed from the same lot of carbide mix and sintered in the same batch. Under these conditions duplicate calibration bars give identical thermoelectric relationships.[6] Otherwise, the tool-work thermoelectric-calibration data when used with the emf of a different tool-work thermocouple may give very misleading results. An extreme case is illustrated in Fig. 1 in which the lower line represents results of a test in which all carbide components are identical and the upper lines are for two random carbide inserts of the same grade and manufacturer, but not of the same thermoelectric characteristics as the "reference" tool used for the lower line. The temperatures for the two random inserts were obtained from the thermoelectric calibration of the reference tool. The authors have found that the tool-chip interface temperature for comparable grades of carbide tools by the several manufacturers is substantially the same, although their thermoelectric powers with respect to a given steel may be distinctly different.

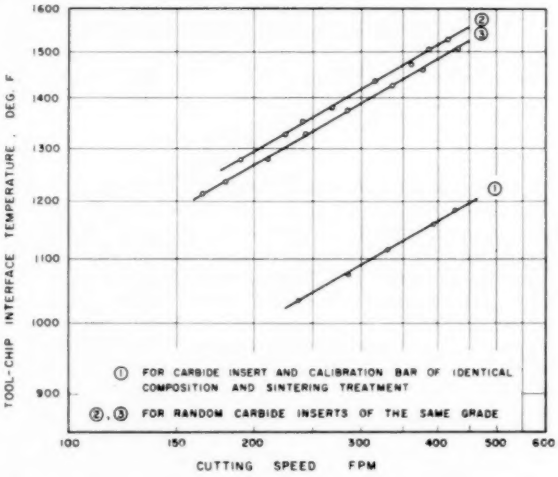


Fig. 1 Effect of Inconsistency in Thermoelectric Characteristics of Calibration Bar and Cutting Inserts on Indicated Tool-Chip Interface Temperature

Working within the limitations imposed by the necessity of identical carbide components the authors have found it most convenient to use a relatively large shank, e.g., 1 in. × 1¼ in. to permit insertion of the carbide contact rod and to provide sufficient tool rigidity to permit a reasonable overhang and thereby facilitate handling of the chip.

However, in order to measure simultaneous values of tool forces and temperature it is usually convenient to use a small (e.g.,

[1] Professor of Mechanical Engineering, University of Illinois, Urbana, Ill. Mem. ASME.

[2] Chicago, Ill.; formerly, Graduate Student, University of Illinois, Urbana, Ill.

[3] Professor of Mechanical Engineering, University of Illinois, Urbana, Ill.

[4] "Temperature Distribution at the Tool-Chip Interface in Metal Cutting," by B. T. Chao and K. J. Trigger, Trans. ASME, vol. 77, 1955, pp. 1107–1121.

[5] "Mechanism of Crater Wear of Cutting Tools," by K. J. Trigger and B. T. Chao, Final Report, Contract DA-11-022-ORD-1121, Office of Ordnance Research, U. S. Army, Durham, N. C., August, 1955.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 6, 1956. Paper No. 56—A-90.

[6] "Progress Report No. 2 on Tool-Chip Interface Temperatures," by K. J. Trigger, Trans. ASME, vol. 71, 1949, pp. 163–174.

1/2-in. square) tool shank to fit the dynamometer. The small shank has insufficient body to accommodate the carbide contact rod for completion of the thermocouple circuit.

The use of a lead, dissimilar in composition to the carbide insert will introduce a parasitic emf, the magnitude and sense of which will depend upon the lead material and the temperature at the point of attachment. The parasitic emf may be sufficient to render results completely misleading.

Czaplicki[7] has reported that Prof. Erich Bickel, ETH, Zurich, Switzerland, devised a method to nullify the parasitic emf caused by leads dissimilar to the tool material. The authors do not have a copy of Bickel's paper nor are they aware of its scope. It is presumed that Bickel's study is not generally known in this country. Consequently, a compensating circuit has been investigated to ascertain its suitability in minimizing the effect of parasitic emf's introduced by the use of dissimilar thermocouple lead materials.

The Compensating Circuit

The need for compensation becomes readily apparent upon examination of Fig. 2, which shows the temperature rise at the rear apex of the small triangularly shaped carbide insert with cutting time. Because of the small mass (~3 1/2 grams) of tool material the trends in Fig. 2 are considered representative of a severe case in so far as temperature rise is concerned.

Ordinarily, a cutting time of 12 to 15 seconds is involved in an interface-temperature measurement (due in part to the time involved in catching the chip and insuring that it is not grounded on the tool shank). In Fig. 2 it is noted that the temperature rise at the rear apex of a steel cutting grade (K3H) is between about 70 and 130 F (~155 to 220 F above zero) depending on the cutting conditions. The temperature rise for the cast-iron cutting grade (K6) is some 250 F in the same time, due not only to its higher thermal conductivity but also since more energy is required for metal removal with such tools on steel. Supplementary tests indicated a steady temperature at the rear apex is approached in 2 to 2 1/2 minutes of cutting.

[7] "Méthodes D'Investigation de L'Usinabilité des Métaux," by L. Czaplicki, *Revue Universelle des Mines*, Imp. De L'Académie, Liège, Belgium, 1952.

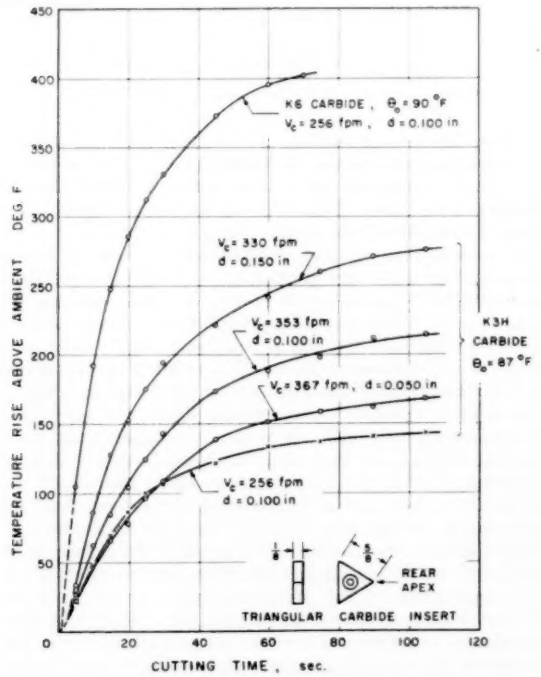


Fig. 2 Temperature Rise at Rear Apex of Small Triangular Carbide Insert
(Work material: leaded 4150 steel, 198 Bhn; tool shape: 0-6-7-7-10-0-0.025 in.)

Considering the temperature rise at the rear apex of the carbide insert it is evident that the use of a dissimilar lead to complete the thermocouple circuit may introduce a serious parasitic emf. The magnitude of this emf was determined with a test setup consisting of a group of common thermocouple wires copper brazed to a carbide calibration bar some 11 in. long, 3/4 in. wide, and 1/8 in. thick. The thermocouple junctions at the braze were thermally insulated to insure a uniform temperature and the assembly placed in an electric furnace. The hot-junction temperature was measured by a standard (Cr Al) thermocouple imbedded in the copper bead and the reference-junction temperature at the external end was maintained by circulating air at ambient temperature.

A precision portable potentiometer was used to measure the thermocouple characteristics of the various combinations. Fig. 3 depicts the temperature-emf characteristics of several common thermocouple metals against K3H carbides and a similar set is shown for K6 carbide in Fig. 4.

An estimate of the error due to the parasitic emf can be obtained from the temperature rise at the rear apex, Fig. 2, and the thermoelectric characteristics shown in Figs. 3 and 4. For example, when cutting dry at ~350 fpm, 0.100 in. depth, and 0.0091 ipr feed, the temperature rise at the point of lead attachment, for K3H carbide, is about 90 F in 15 seconds. If the least active lead (alumel or nickel) were used, the parasitic emf amounts to about 0.3 mv, negative with respect to the carbide and boosting the emf generated at the tool-chip interface. The millivolt power of a steel-cutting grade of carbide with steel is of the order of 0.01 mv/deg F and therefore the indicated tool-chip interface temperature would be approximately 30 F high. The most active lead, chromel, would similarly introduce a parasitic emf of 1.7 mv of opposing polarity and result in an indicated temperature some 170 F low.

In Fig. 4 it is noted that the thermoelectric effect of alumel and K6 carbide is negligible for a temperature rise of about 150 F. If one were to use this combination with a cutting fluid in sufficient quantity to maintain the lead-junction temperature near ambient, no significant error would be introduced. As an extreme case in dry cutting, the use of a chromel lead with a temperature rise of about 250 F in 15 seconds would introduce a parasitic emf of about 6 mv and result in an indicated tool-chip interface temperature some 300 F low.

A compensating circuit to minimize the parasitic emf's is illustrated schematically in Fig. 5 with the usual polarity as indicated. The principal thermocouple is 1–2 at the tool work interface and the parasitic emf's are introduced at 2–3 and 4–2, the junctions of the carbide with the thermocouple members of the compensating circuit. Fixed resistances R_A and R_B provide an IR drop in the closed compensating circuit. The lead to the potentiometer is attached at point Y such that the IR drop in each leg of the circuit, $i(R_3 + R_A)$ or $i(R_4 + R_B)$ is theoretically equal to and opposite in sense to the parasitic emf generated in the respective carbide-lead pair (2–3 or 4–2) by the temperature rise at point X. Complete compensation is attained only when such a balance exists.

It is apparent that the ratio of the resistances in each leg of the compensating thermocouple circuit must be the same as the ratio of the respective thermoelectric effects with the carbide insert if

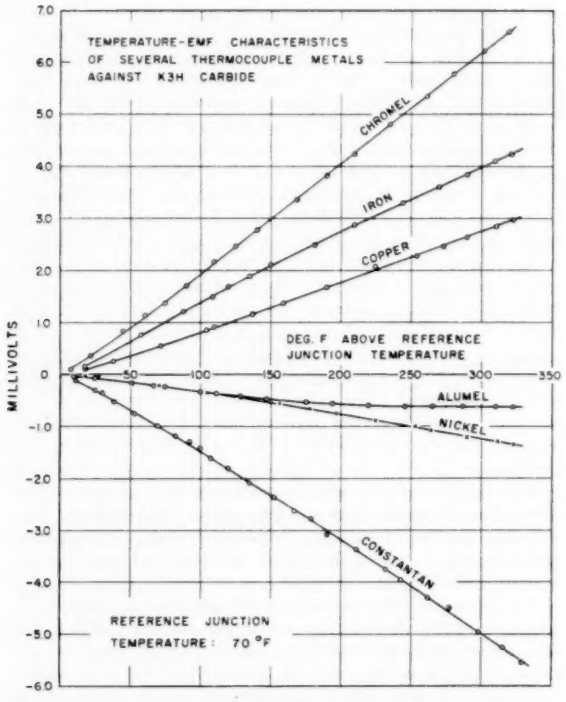


FIG. 3 THERMOELECTRIC CHARACTERISTICS OF COMMON THERMOCOUPLE METALS AGAINST K3H CARBIDE

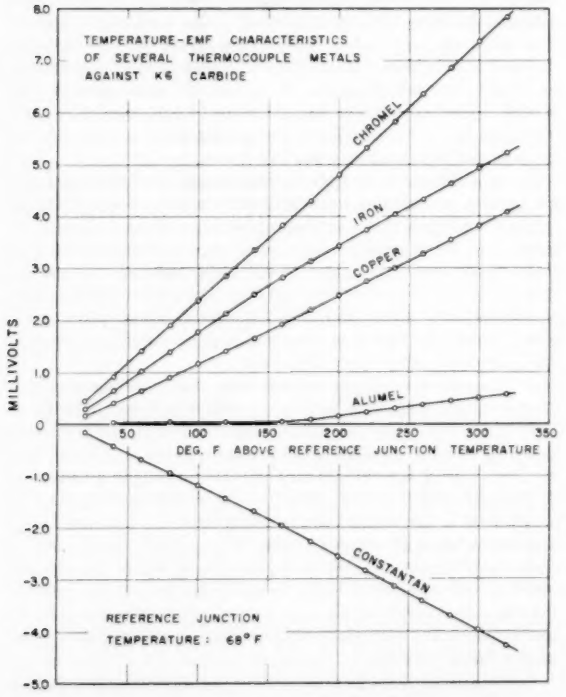
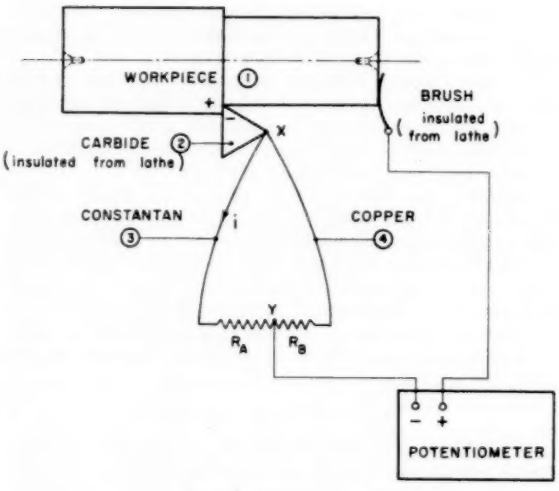
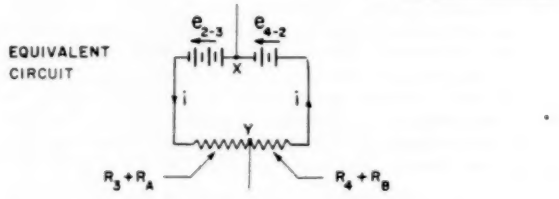


FIG. 4 THERMOELECTRIC CHARACTERISTICS OF COMMON THERMOCOUPLE METALS AGAINST K6 CARBIDE





FOR COMPLETE COMPENSATION:

$$e_{2-3} = i\,(R_3 + R_A)$$

$$e_{4-2} = i\,(R_4 + R_B)$$

FIG. 5 COMPENSATING CIRCUIT

compensation is to be realized. It is equally apparent that the compensating leads must be of opposite polarity with respect to the carbide if compensation is to be attained. Otherwise, the minimum error is that of the lead with the lowest thermoelectric power against the carbide insert.

As a consequence of these observations two criteria for the compensating leads are presented: (*a*) The thermoelectric characteristics of the leads with the tool material must have the same kind of temperature dependence (linear is ideal), and (*b*) the leads must have opposite polarity with respect to carbide. Attention was centered on the steel-cutting grade (K3H) carbide and the several possibilities suggested by Fig. 3 were explored.

In order to facilitate tool sharpening, light (~24-gage) compensating leads were used. Four-step, decade-resistance boxes were used for R_A and R_B of the compensating circuit. A total resistance of 50 to 100 ohms in the compensating circuit was selected to minimize the effect of changes in the resistance of the leads due to occasional breakage in use. The total resistance of the tool-work thermocouple circuit was in the vicinity of 20 ohms, well within the working range of the potentiometers in use.

An examination of Fig. 3 suggests that the most promising negative leads are A-nickel and constantan. While both were investigated attention was directed toward the latter since it is a standard thermocouple element. Considering the temperature dependence of its thermoelectric characteristics with K3H carbide, the positive lead most nearly like constantan is copper. Chromel appears to be the second choice and iron the third. Other positive thermocouple leads such as platinum and platinum-

10 per cent rhodium were investigated but are unfavorable cost-wise and are not included in this paper.

The effectiveness of the compensating circuit for a given pair of leads was evaluated by fixing resistance R_A and varying R_B to provide complete compensation as indicated by zero unbalance on the precision potentiometer. One potentiometer lead was attached to the carbide calibration bar and the other to point Y, Fig. 5. Once the range of R_B was found a value was selected to provide complete compensation at some point in the expected temperature range; e.g., 200 F.

A series of tests also was conducted with the K6 carbide. A comparison of the thermoelectric characteristics of the two grades, Figs. 3 and 4, indicates differences in the individual effects with carbide. While the algebraic sum of the thermoelectric emf of a pair of thermocouple wires with carbide is independent of the carbide material used (by virtue of the law of intermediate metals) it is the individual emf with respect to carbide which governs the resistances R_A and R_B in the compensating circuit. A change of tool material will (usually) affect the magnitude of R_A and R_B. As can be seen in Fig. 4 the only negative lead (of those studied) is constantan, and, again, copper is the most promising positive lead.

The unbalance at the potentiometer for several compensating-circuit combinations nulled at ~200 F is shown in Fig. 6, and the resistance in each leg is given in Table 1.

TABLE 1 RESISTANCE IN EACH LEG OF COMPENSATING THERMOCOUPLE CIRCUIT SUMMARIZED IN FIG. 6

Compensating couple	Resistance, ohms With K3H carbide	Resistance, ohms With K6 carbide
Copper	29.57	51.97
Constantan	54.26	54.26
Iron	59.92	...
Constantan	54.26	...
Chromel	54.37	...
Alumel	8.14	...

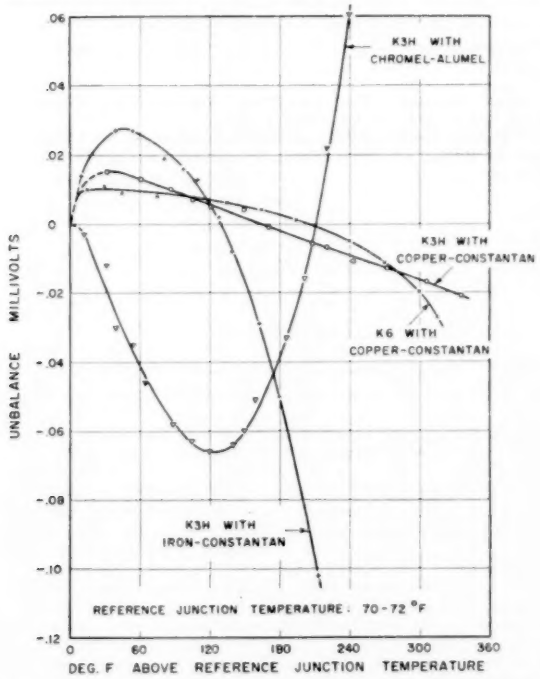


FIG. 6 UNBALANCE IN SEVERAL COMPENSATING CIRCUITS OVER A RANGE OF TEMPERATURE

It is apparent that the most satisfactory pair over the range illustrated (temperature rise at 315 F or a rear apex temperature of ~385 F) is copper constantan, an inexpensive, readily available thermocouple material. The unbalance of about ±0.02 mv for copper-constantan corresponds to ∓2 F for steel-cutting grades of carbide. Obviously, such an error may be ignored completely. Indeed, any of the combinations illustrated would introduce a maximum error of about 5 F if used in a cutting temperature test of 15 seconds' duration.

A cutting tool was equipped with both a carbide contact rod for use in the conventional manner and a copper-constantan compensating circuit attached at the rear of the brazed insert.

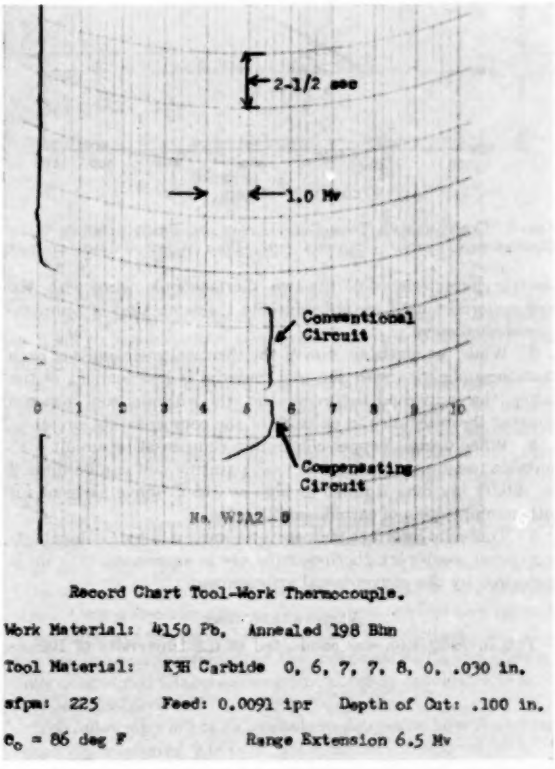


FIG. 7 RECORD CHART OF CONVENTIONAL AND COMPENSATING TOOL-WORK THERMOCOUPLE

Either tool-work circuit could be engaged during a cut and a typical record chart of a test is shown in Fig. 7. Clearly, both circuits indicate the same millivolt reading. A series of cutting tests using both circuits is shown in Fig. 8 in which the reliability of the compensating circuit is clearly demonstrated.

CONCLUSIONS

1 The temperature at the rear apex of a small triangularly shaped carbide insert attains a value sufficiently high to introduce a serious parasitic emf if a dissimilar conductor is used to connect the carbide insert to the emf-sensing device in a tool-work thermocouple.

2 The effect of the parasitic emf can be minimized by the use of a compensating thermocouple circuit. Two criteria are indicated: (*a*) The thermocouple wires must have opposite polarity with respect to the tool material, and (*b*) the thermo-

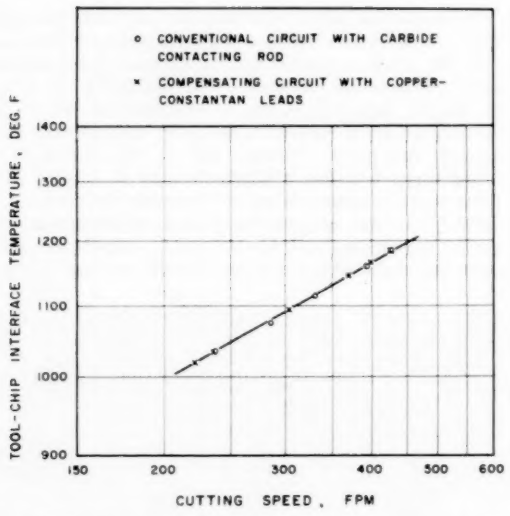


FIG. 8 COMPARATIVE TOOL-CHIP INTERFACE TEMPERATURES WITH CONVENTIONAL AND COMPENSATING TOOL-WORK THERMOCOUPLE

electric characteristics of the two thermocouple wires with the tool material must have substantially the same kind of temperature dependence.

3 While the algebraic sum of the thermoelectric emf's of each thermocouple wire with the tool material is independent of the latter, the individual thermoelectric emf with the tool material governs the resistance in each leg of the compensating circuit.

4 When using a copper-constantan compensating circuit with suitable resistances, the unbalanced parasitic emf can be limited to ± 0.02 mv over a range as high as 315 F above ambient for either composition of carbide investigated.

5 Tool-chip interface temperatures measured by a compensating-circuit, tool-work thermocouple are in agreement with those measured by the conventional arrangement.

ACKNOWLEDGMENT

This investigation was conducted at the University of Illinois as part of a program sponsored by the Office of Ordnance Research, U.S. Army, under Contract No. DA-11-022-ORD-1980. The authors hereby express their sincere appreciation to that Office for the support of this program. Acknowledgment is made to Kennametal, Inc., Latrobe, Pa., and to Mr. W. L. Kennicott, for the carbide tool materials used in the investigation, and to Miss Irene Cunningham for the typing of the manuscript.

Discussion

E. G. LOEWEN.[8] The authors have given a very complete description of how to get the most out of the compensating circuits that get around one of the difficulties in tool-work thermocouple measurements. The paper will be of obvious help to all workers in this field.

To complete the record, Professor Bickel's original publication of this idea was published in 1950.[9]

The curves in Figs. 3 and 4 show what the writer discovered some years ago, namely, that 90 per cent of the time one can stay well within normal measuring errors by making connection to the carbide-tool bit with a single alumel wire. Only in special cases will the superior copper-constantan method be required.

AUTHORS' CLOSURE

The authors appreciate Dr. Loewen's comments and his reference to the original paper. However, they call attention to the fact that an error of 30 to 50 deg F may be very important, particularly in the temperature sensitive range of carbide tools. At higher cutting speeds (and temperatures) not only is the error due to parasitic emf greater, but the consequences of that error are multiplied.

The authors were concerned with clarification of the conditions necessary for complete compensation as well as the best compensating pair over a temperature range. That copper constantan is most suitable is evident in Fig. 6 of the paper. Its use involves no complications and the error may be completely ignored.

---

[8] Staff Engineer, The Taft-Peirce Manufacturing Company, Woonsocket, R. I. Assoc. Mem. ASME.

[9] "Die Zerspannungsforschung am Werkzeugmaschinen-Laboratorium der ETH," by Erich Bickel, *Industrielle Organisation*, No. 4, 1950, pp. 1–7.

On the Theoretical Analysis of a Dynamic Thermocouple[1]

By E. W. GAYLORD,[2] W. F. HUGHES,[3] F. C. APPL,[3] and F. F. LING,[4] PITTSBURGH, PA.

The "dynamic" thermocouple formed by the moving junctions of two dissimilar metals is analyzed theoretically. It is found that if the two leads from the cold junction, in series with a potentiometer, are symmetrically placed in two bodies rubbing over each other, the e.m.f. measured by the potentiometer satisfies Laplace's equation in terms of the positioning of the leads in the body. The boundary condition is that the potential at any contact area is the Seebeck e.m.f. corresponding to the contact area temperature. It is shown that, in the case of two semi-infinite rubbing bodies with many randomly distributed contacts, small in area compared to the distance between them, the potential measured by thermocouple leads placed at an infinite distance away from the contact areas is the average of the Seebeck e.m.f.'s, corresponding to the contact temperatures, weighted by the square root of the areas.

Nomenclature

The following nomenclature is used in the paper:

E = Seebeck e.m.f.
π_{AB} = Peltier coefficient
σ = Thomson coefficient
γ = electrical resistivity
J = current flux density
P = potential measured by a potentiometer placed in the thermocouple circuit
ϕ = $P - P(\infty)$
$\overrightarrow{m_n(r)}$ = intensity of the current source for the nth contact area
(x, y, z) = co-ordinates in body B
(ξ, η, ζ) = co-ordinates in body A
$T(x, y, z)$ = temperature in body B
T_0 = cold junction temperature
$(\ )_A$ = subscript referring to body A
$(\ )_B$ = subscript referring to body B

Introduction

The dynamic thermocouple, better known as the "Herbert Gottwein" thermocouple utilizes the junction between two dissimilar metals which are rubbing together as one junction of a thermocouple circuit, Fig. 1. The e.m.f. generated in this circuit is used to estimate the temperature at the interface of two rubbing dissimilar metals.

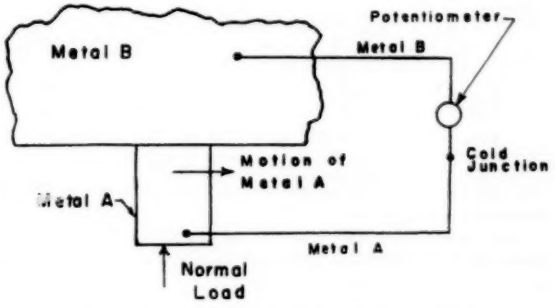


Fig. 1 Schematic of Dynamic Thermocouple

In metal cutting research the dynamic thermocouple, formed by the cutting tool and its moving tool chip, has been used to estimate the average interface temperature. In this application, the measured interface temperature has been found to agree rather well with the theoretically determined area-averaged interface temperature (1).[5]

The dynamic thermocouple has also been applied to the investigation of frictional heating between two metallic surfaces sliding over each other (2). In this case the force between the two rubbing surfaces is much lower than it is between a tool and tool chip. Only a small fraction of the apparent contact area between the interface of the two rubbing metals is true contact area. Correlation between the measured thermocouple e.m.f. and theoretically determined temperatures has not been as conclusive as in the tool chip problem.

The following work is an analysis to determine how the instantaneously measured e.m.f. of the dynamic thermocouple is related to the instantaneous temperature distribution over the contact interface of the moving junction.

Analysis

Consider as a model a large body of metal B rubbing on a large body of metal A at a given instant of time with axes (x, y, z) and (ξ, η, ζ) respectively; see Fig. 2. With a lead of metal B located at (x, y, z) in body A, the potential measured by the potentiometer will be some function $P(x, y, z, \xi, \eta, \zeta)$.

It will be assumed that the rubbing does not in itself generate any voltages in the system so that the potential is the same as it would be in a static system with the same temperature and contact situation. Assume that electrical time constants due to inductive and capacitive effects are zero so that electrically the system has reached steady state.

The interface between the two metals may be considered to consist of any number of arbitrarily small contact areas which in

[1] This work was done in part under the sponsorship of Watertown Arsenal Laboratory, Watertown, Mass., under Contract No. DA-36-061-ORD-400.

[2] Assistant Professor, Department of Mechanical Engineering, Carnegie Institute of Technology. Assoc. Mem. ASME.

[3] Research Assistant, Department of Mechanical Engineering, Carnegie Institute of Technology.

[4] Assistant Professor of Mathematics, Department of Mathematics, Carnegie Institute of Technology. Assoc. Mem. ASME.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 6 1956. Paper No. 56—A-86.

[5] Numbers in parentheses refer to the Bibliography at the end of the paper.

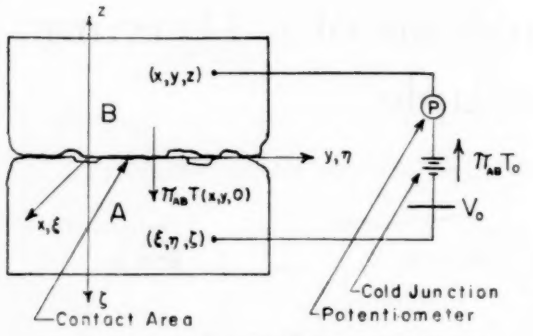


FIG. 2 MODEL FOR DYNAMIC THERMOCOUPLE CIRCUIT

the limit become a continuous contact area with any arbitrary temperature distribution.

Next consider the thermoelectric effects in the thermocouple. Referring to an ordinary thermocouple, see Fig. 3, the Seebeck e.m.f. of a thermocouple is

$$E = \pi_{AB}T_2 - \pi_{AB}T_1 + \int_{T_1}^{T_2} (\sigma_A - \sigma_B)dT \quad [1]$$

For the purpose of analysis the Seebeck e.m.f. is divided into two parts; the Peltier e.m.f., $\pi_{AB}T_1 - \pi_{AB}T_2$ which will be assumed to be generated at the junctions of the dissimilar metals, and the Thomson e.m.f., $\int_{T_1}^{T_2} (\sigma_A - \sigma_B)dT$ which is assumed to be generated in the metal.

Let us now consider relations which must hold within the metal *B*. First consider the one dimensional case, Fig. 4. Taking account of the Thomson e.m.f. and the Joule effect in the element

$$V_{x+\Delta x} - V_x = \int_{T_x}^{T_{x+\Delta x}} \sigma_B\, dt - J\frac{\Delta x}{\gamma} \quad [2]$$

from which

$$J_{(x)} = \gamma\sigma_B \frac{dT}{dx} - \gamma \frac{dV}{dx} \quad [3]$$

By similar reasoning for three dimensions

$$J = \gamma\sigma_B\nabla_B T - \gamma\nabla_B V \quad [4]$$

For steady state current flow

$$\nabla \cdot J = 0 \quad [5]$$

Combine [4] and [5]

$$\nabla_B{}^2 V = \nabla_B \cdot (\sigma_B \nabla_B T) \quad [6]$$

Referring to Fig. 2 note that

$$V_{(x,y,z)} = P + \int_{T_0}^{T(x,y,z)} \sigma_B dT + V_0 - \pi_{AB}T_0 \quad [7]$$

From [7] one finds that

$$\nabla_B{}^2 V = \nabla_B{}^2 P + \nabla_B(\sigma_B \nabla_B T) \quad [8]$$

Combining [6] and [8]

$$\nabla_B{}^2 P = 0 \quad [9]$$

By similar reasoning

$$\Delta_A{}^2 P = 0 \quad [10]$$

Along the boundaries with no current flow $J_n = 0$

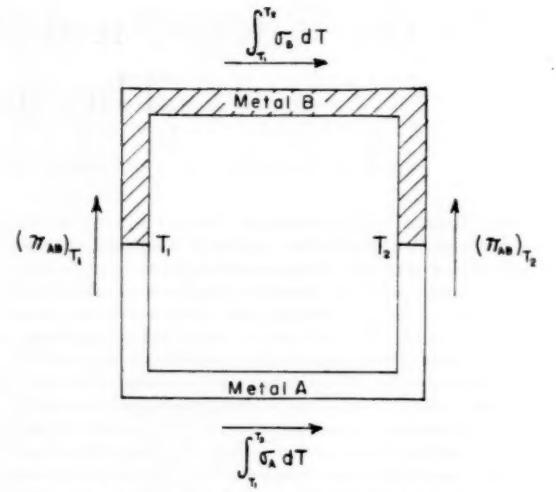


FIG. 3 ILLUSTRATION OF THERMOCOUPLE EFFECTS

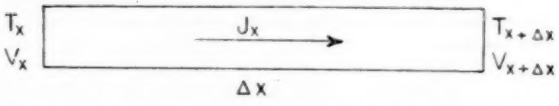


FIG. 4 ONE DIMENSIONAL CURRENT FLOW IN THE METALLIC SOLID

Equation [4] becomes

$$\sigma_B\nabla_B T - \nabla_B V = 0 \quad [11]$$

Combining [11] and [7] gives

$$\frac{\partial P}{\partial n} = 0 \quad [12]$$

For the boundary conditions where there is contact, the voltage drops around the circuit equal zero

$$P = \int_{T_0}^{T} (\sigma_B - \sigma_A)dT + \pi_{AB}T - \pi_{AB}T_0 - B\int_{x,y,0}^{x,y,z} \frac{J}{\gamma}\cdot dS - A\int_{\xi,\eta,0}^{\xi,\eta,\zeta} \frac{J}{\gamma}\cdot dS \quad [13]$$

or

$$P = E(T_{x,y,0}) - B\int_{x,y,0}^{x,y,z} \frac{J}{\gamma}\cdot dS - A\int_{\xi,\eta,0}^{\xi,\eta,\zeta} \frac{J}{\gamma}\cdot dS \quad [14]$$

where $E(T_{x,y,0})$ is the Seebeck e.m.f. corresponding to the temperature on the boundary.

Since this boundary condition involves both blocks, *A* and *B*, the potential is a function $P(x, y, z, \xi, \eta, \zeta)$, and Equations [9] and [10] would have to be solved simultaneously.

This problem may be simplified by assuming that both blocks are identical with regard to boundary conditions and that the leads are symmetrically placed at $(x, y, z) = (\xi, \eta, \zeta)$.

This assumption would be reasonable if both blocks were semi-infinite and the leads were placed an infinite distance away from the contact area. With this assumption

$$\nabla_A = \nabla_B = \nabla \quad [15]$$

By simultaneously placing both leads at $(x, y, 0)$ and $(\xi, \eta, 0)$,

Equation [14] gives the boundary condition

$$P = E(T_{x,y,z}) \quad \ldots\ldots\ldots\ldots\ldots\ldots [16]$$

Hence we now have the problem reduced mathematically to a potential problem in a model the shape of one of the rubbing bodies, with the boundary conditions that the potential is equal to the Seebeck e.m.f. on the boundary where rubbing contact is made, and $\frac{\partial P}{\partial n} = 0$ on the boundary where no contact is made.

In most practical applications of the dynamic thermocouple where the contact area is small compared to the size of the rubbing bodies one would be interested in solving for the potential at infinity in terms of the contact area temperatures and their corresponding Seebeck e.m.f.'s.

Many Randomly Distributed Small Contact Areas

The foregoing theory will be applied to a case which might represent the frictional rubbing of two metal surfaces under a light load, where it will be assumed that there are many contact asperities and the total real contact area is small compared to the apparent contact area. The problem is to determine what temperature a potentiometer would indicate if the thermocouple leads are placed a large distance away from the contact areas.

In reducing this case to a potential problem as previously shown, consider a semi-infinite solid, Fig. 5. Near the origin on the surface there is a distribution of area sources of current. These sources are held at constant voltage, each source being at some value. The size and value of the potential of each source vary.

It is desired to find the potential in the solid at some large distance removed from the localized region of sources. With $P(\infty)$ defined as the potential at this point, define for convenience $\phi = P - P(\infty)$.

As previously derived, $\nabla^2\phi = 0$, $\frac{\partial \phi}{\partial n} = 0$ on surfaces with no contact and $\phi = E_n - P(\infty) = V_n$ on surface of the nth source area of a total of N source areas.

Contributions to $\phi(x, y, z)$ due to each source area can be added. Let $\vec{R}$ denote the point where potential is to be measured and $\vec{r}$ a point in the area source, then the contribution of the nth source to the potential is

$$\phi_n(\vec{R}) = \int_{A_n} \frac{m_n(\vec{r}_n)\,dA_n}{|\vec{R}_n - \vec{r}_n|} \quad \ldots\ldots\ldots\ldots [17]$$

where $m_n(\vec{r}_n)$ is the intensity of the source as a function of $\vec{r}_n$.

For N sources

$$\phi(\vec{R}) = \sum_{n=1}^{N} \int \frac{m_n(\vec{r}_n)}{|\vec{R}_n - \vec{r}_n|}\, dA_n \quad \ldots\ldots\ldots\ldots [18]$$

and on the ith source Equation [18] expresses the boundary condition as

$$V_i = \sum_{n=1}^{N} \int_{A_n} \frac{m_n(\vec{r}_n)}{|\vec{R}_i - \vec{r}_n|}\, dA_n \quad \ldots\ldots\ldots\ldots [19]$$

For large values of R_n, Equation [18] makes $\phi(\vec{R})$ go to zero for finite sources, which is consistent with the definition of ϕ.

The solution of the problem lies in finding the value of the current intensity $m_n(\vec{r}_n)$ in terms of the source potentials V_n. To simplify this step divide Equation [19] into two parts

$$V_i = \sum_{\substack{n=1 \\ n \neq i}}^{N} \int_{A_n} \frac{m_n(\vec{r}_n)}{|\vec{R}_i - \vec{r}_n|}\, dA_n + \int_{A_i} \frac{m_i(\vec{r}_i)}{|\vec{R}_i - \vec{r}_i|}\, dA_i \quad \ldots [20]$$

Examine the first part of [20]. If the areas A_n are small compared to values of R_i one could write

$$\sum_{\substack{n=1 \\ n \neq i}}^{N} \int_{A_n} \frac{m_n(\vec{r}_n)}{|\vec{R}_i - \vec{r}_n|}\, dA_n = \sum_{\substack{n=1 \\ n \neq i}}^{N} \frac{\int_{A_n} m_n(\vec{r}_n)\,dA_n}{\overline{|\vec{R}_i - \vec{r}_n|}} \quad \ldots [21]$$

where $\overline{|\vec{R}_i - \vec{r}_n|}$ is a mean value.

From continuity of currents

$$\sum_{n=1}^{N} \int_{A_n} m_n(\vec{r})\,dA_n = 0 \quad \ldots\ldots\ldots\ldots [22]$$

On the basis of Equation [22], if there is a random distribution of sources with respect to current flux, $\int_{A_n} m_n(\vec{r}_n)dA_n$, and if all values of $|\vec{R}_i - \vec{r}_n|$ are large compared to $|\vec{R}_i - \vec{r}_i|$, then the first part of Equation [20] will go to zero as N, the number of sources becomes very large, giving

$$V_i = \int_{A_i} \frac{m_i(\vec{r}_i)}{|\vec{R}_i - \vec{r}_i|}\, dA_i \quad \ldots\ldots\ldots\ldots [23]$$

To evaluate Equation [17] assume that all sources are circles of radius a. The current strength for such a source is given in reference (4)

$$m_i(\beta) = \frac{K_i}{\sqrt{(a_i^2 - \beta_i^2)}} \quad \ldots\ldots\ldots\ldots [24]$$

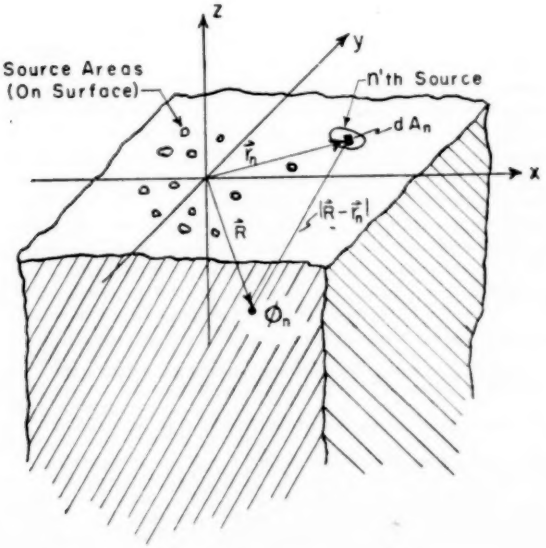


Fig. 5 Randomly Distributed Current Sources

where $\beta_i = |\vec{R}_i - \vec{r}_i|$ and K is some proportionality constant to be evaluated. Substituting [24] in [23] yields $K_i = \frac{V_i}{\pi^2}$ giving

$$m_i(\beta_i) = \frac{V_i}{\pi^2 \sqrt{(a_i^2 - \beta_i^2)}} \quad [25]$$

Substituting [25] in [22]

$$0 = \sum_{n=1}^{N} \int_{A_n} m_n(\beta_n) dA_n = \Sigma V_n a_n \quad [26]$$

$$A_n = \pi a_n^2; \quad \text{hence } \Sigma V_n \sqrt{A_n} = 0 \quad [27]$$

Since

$$V_n = E_n - P(\infty)$$

$$P(\infty) = \frac{\Sigma E_n \sqrt{A_n}}{\Sigma \sqrt{A_n}} \quad [28]$$

Equation [28] gives the conclusion that, if the Seebeck e.m.f. is linearly related to the temperature, one would measure the square root area average temperature rather than the area average temperature.

In the friction problem the physical interpretation would be that, if there were any correlation between the size of the protuberances and their temperature, then the temperature corresponding to the dynamic thermocouple e.m.f. would be weighted more heavily to the temperature of the smaller protuberances.

BIBLIOGRAPHY

1 "Temperature Distribution at the Tool-Chip Interface in Metal Cutting," by B. T. Chao and K. J. Trigger, Trans. ASME, vol. 77, 1955, pp. 1107–1121.

2 "The Friction and Lubrication of Solids," by F. P. Bowden and D. Tabor, Oxford University Press, London, England, 1950.

3 "Electricity and Magnetism," by F. W. Sears, Addison-Wesley Publishing Company, Cambridge, Mass., 1953, pp. 160–168.

4 "Foundations of Potential Theory," by O. D. Kellog, Dover Press, New York, N. Y., 1953, p. 188.

Temperature Distribution at Tool-Chip and Tool-Work Interface in Metal Cutting

By B. T. CHAO[1] and K. J. TRIGGER,[2] URBANA, ILLINOIS

A noniterative method is presented for the computation of temperature distribution both at the tool-chip and tool-work interface in metal cutting. Temperatures at the tool-work interface increase appreciably with the increase in flank wear. This phenomenon contrasts to the relatively minor influence on tool-chip interface temperature as crater wear develops. Results include a three-dimensional temperature distribution at the tool top surface.

Introduction

THIS paper is based upon two Technical Reports (1, 2)[3] issued in pursuit of contracted research for the Office of Ordnance Research, U. S. Army. Some modifications and additions have been made and this report represents an extension of an earlier work (3) developed by the authors on the analytical evaluation of tool chip interface temperature distribution. The use of the point source equation results in an increasing error when the number of network subdivisions exceeds a certain limit. This difficulty is removed by replacing it with an area source equation. The general method of approach remains essentially unaltered but following a suggestion by Blok (4) the procedure has been made noniterative. The effect of heat generation due to rubbing at the tool flank is included in the present analysis. Nomenclature used in the earlier publications is retained wherever feasible.

Basic Assumptions

This analysis is concerned with the corner cutting by an orthogonal tool producing a type 2 chip. The three distinct regions of heat generation considered are: (*a*) the shear zone, *OW*, where the main plastic deformation takes place, (*b*) the tool-chip interface, *OT*, where the heated chip slides on the tool top surface, and (*c*) the tool-work interface, *OF*, where frictional rubbing of the workpiece on the tool flank occurs. This is shown in Fig. 1.

Several assumptions are used in the present analysis, to wit:

(*a*) The heat flow is steady in the cutting tool and quasi-steady in the moving chip and workpiece. Average interface temperature records taken during turning operations indicate the validity of this assumption. On the other hand, intermittent machining operations like milling produce unsteady heat flow.

(*b*) All of the mechanical work of plastic deformation is converted into sensible heat. Actually, a small fraction of such work of deformation is retained as latent energy in the strain-hardened chip and is thus not available to raise its temperature. Discussions on the possible magnitude of error introduced as a consequence of this assumption have been given (3).

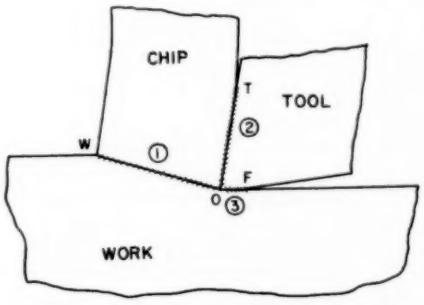
(*c*) The distribution of shear energy in region ① and frictional energy in regions ② and ③ are uniform. This assumption was adopted because of lack of information on the subject. It is pertinent that with the method of calculation herein presented, a nonuniform liberation of heat at both interfaces introduces no complication.

(*d*) The chip as it is formed at the shear zone has a uniform temperature. While the problem of temperature distribution along the shear plane has been treated in the literature (5, 6), the uncertainties involved in the analyses, particularly with respect to the local temperature of the chip in the close vicinity of the cutting edge (point *O* in Fig. 1), do not warrant their inclusion at the present time. At large values of thermal number $\left(R_t = \frac{v_c d}{\kappa}\right)$,[4] this assumption is a close approximation (5).

(*e*) The dimensions of the tool are large compared to the cutting geometry and it can be considered as infinite in extent insofar as the temperature rise over the two interfacial areas is concerned. Corner radius, rake, and clearance angles of the tool are all taken as zero. Such an idealization greatly simplifies the mathematics involved. Experience in the solution of this problem over practical ranges of cutting conditions indicates that this assumption would not introduce serious error. Steps are now being taken to evaluate quantitatively the combined influence of rake and clearance angles on the conducting capacity of the tool.

(*f*) Heat loss at all surfaces of the tool and that at the chip and workpiece surfaces are ignored. This assumption is, in general, valid since the heat flux at the two rubbing contacts is usually many thousand-fold that at the exposed tool surfaces. Likewise, due to the relatively high sliding velocity at both the chip and work surface in contact with the tool, the quasi-steady tempera-

[4] The thermal number is defined as the ratio of the product of cutting speed and feed to the thermal diffusivity of the work material.



① DUE TO MAIN CHIP SHEAR

② DUE TO FRICTIONAL RUBBING AT THE TOOL TOP SURFACE

③ DUE TO FRICTIONAL RUBBING AT THE TOOL FLANK

Fig. 1 Heat Sources in Metal Machining—Orthogonal Tool and Type 2 Chip

[1] Professor of Mechanical Engineering, University of Illinois.

[2] Professor of Mechanical Engineering, University of Illinois. Mem. ASME.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 6, 1956. Paper No. 56—A-87.

ture distribution is practically unaffected by the surface heat loss.

(*g*) In the calculation of average chip bulk temperature, the variation of thermal properties of the work material with temperature changes may be properly taken into consideration in a manner described in reference (3). Nevertheless, opinions still differ regarding the temperature at which the thermal conductivity and specific heat of the chip and work material should be evaluated for the computation of temperatures at the two sliding contacts. In this paper, these properties are taken at the bulk temperature of the chip for the tool-chip interface and at the bulk temperature of the workpiece for the tool-work interface. This was adopted following the consideration of the extremely steep temperature gradient at the rubbing surfaces of the chip and workpiece. Loewen and Shaw (7) favor the use of average interface temperature. At present, it is not known which procedure yields a better result. A value intermediate between the bulk and average interface temperature would probably be closer to the facts. This latter procedure, as well as that of Loewen and Shaw, entails iteration.

(*h*) The thermal conductivity of the steel cutting-grade carbide is unaffected by temperature changes This assumption is well justified in view of the recently published data by Loewen (8).

(*i*) The temperature distribution at both interfaces is uniform in a direction normal to chip motion, i.e., parallel to the active cutting edge of the orthogonal tool. The chip is regarded as a semi-infinite solid with band source of *variable* intensity in the direction of its motion relative to the tool, and a similar assumption is made for the workpiece. A detailed three-dimensional analysis indicates the general suitability of this idealization for the present purpose. Results obtained from the latter analysis are given at the end of the paper.

Outline of the Noniterative Method of Computation and the Accompanying Equations

Fig. 2 shows the cross section of a worn tool illustrating the moving heat sources at the chip and work surfaces in the ideal case. The tool-chip interface is subdivided into m bands of width $2\Delta x$, and the tool-work interface is subdivided into n bands of width $2\Delta y$. Clearly, $\Delta x = \frac{l_c}{m}$, $\Delta y = \frac{l_w}{n}$. Considered from the point of view of the chip, one then has m contiguous band sources of heat, moving at the chip flow velocity v_f, but in a direction opposite to the actual chip motion relative to the tool. The quasi-steady temperature rise above $\bar{\theta}_c$, the average chip bulk temperature, at the center of any such subarea i' due to a uniform heat flux $q_{c,i}$ over the band source i is, according to Jaeger (9)

$$\delta\theta_{c,i'i} = \frac{2q_{c,i}}{\pi C_c \rho v_f} \int_{E_{ii'}-\Delta X}^{E_{ii'}+\Delta X} e^{-u} K_0(|u|)du \quad \ldots\ldots [1]$$

where $E_{ii'} = \frac{v_f \xi_{ii'}}{2\kappa_c}$, $\Delta X = \frac{v_f \Delta x}{2\kappa_c}$. κ_c, C_c, and ρ are respectively the thermal diffusivity, specific heat, and density of the chip material evaluated at $\bar{\theta}_c$. $\xi_{ii'}$ is the distance measured from i to i', being positive in a direction opposite to that of chip motion as shown in Fig. 2. K_0 is the modified Bessel function of the second kind and zero order and $E_{ii'}$ and ΔX are dimensionless.

The definite integral which appears in Equation [1] cannot be expressed in closed form. To facilitate its numerical evaluation, one writes

$$\int_A^B e^{-u} K_0(|u|)du = I(B) - I(A)$$

with

$$I(p) = \int_0^p e^{-u} K_0(|u|)du$$

and

$$I(-p) = -\int_0^p e^{u} K_0(|u|)du$$

The values of $I(p)$ and $I(-p)$ over a range of p values from -100

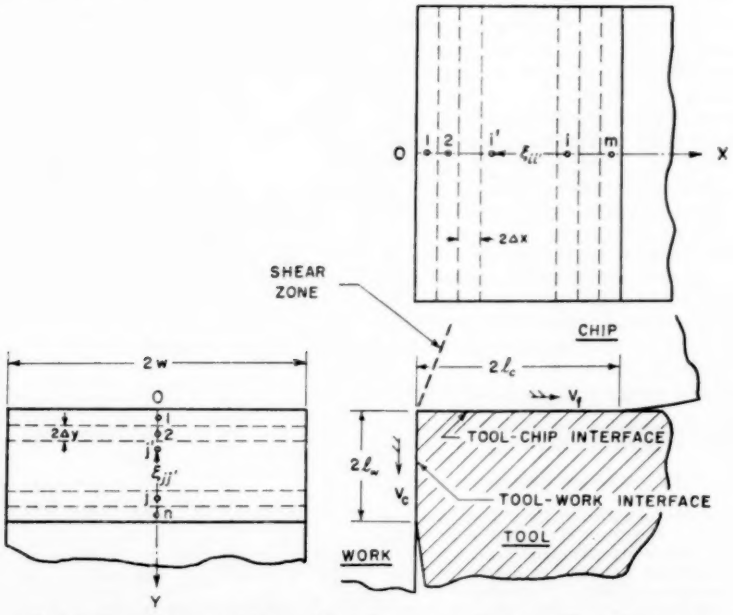


Fig. 2 Subdivision of Moving Heat Sources at Chip and Work Surfaces

to large positive are shown graphically in Fig. 3. Having selected the number of sub-divisions m, Δx becomes fixed in a given problem. The value of the previous definite integral depends solely on the dimensionless position parameter $E_{ii'}$. For convenience, one designates

$$\int_{E_{ii'}-\Delta X}^{E_{ii'}+\Delta X} e^{-u}K_0(|u|)du = {}_iI_{i'} \quad \ldots\ldots\ldots\ldots [2]$$

It follows that the resultant temperature at i' due to m such band sources is

$$\theta_{w,i'} = \bar{\theta}_0 + \frac{2}{\pi C_w \rho v_f} \sum_{i=1}^{m} q_{w,i}({}_iI_{i'}) \quad \ldots\ldots\ldots\ldots [3]$$

Likewise, at the work-flank interface, one notes that the quasi-steady temperature rise above θ_0, the initial uniform temperature of the workpiece, at the center of any subarea j' due to a uniform heat flux $q_{w,j}$ over the band source j is

$$\Delta\theta_{w,j'j} = \frac{2q_{w,j}}{\pi C_w \rho v_c} \int_{E_{jj'}-\Delta Y}^{E_{jj'}+\Delta Y} e^{-u}K_0(|u|)du \quad \ldots\ldots [4]$$

where $E_{jj'} = \dfrac{v_c \xi_{jj'}}{2\kappa_w}$, $\Delta Y = \dfrac{v_c \Delta y}{2\kappa_w}$. v_c is the cutting speed, and κ_w and C_w are to be evaluated at θ_0. The resultant temperature at j' due to n such band sources is

$$\theta_{w,j'} = \theta_0 + \frac{2}{\pi C_w \rho v_c} \sum_{j=1}^{n} q_{w,j}({}_jI_{j'}) \quad \ldots\ldots\ldots\ldots [5]$$

with

$${}_iI_{i'} = \int_{E_{ii'}-\Delta Y}^{E_{ii'}+\Delta Y} e^{-u}K_0(|u|)du$$

Considered from the point of view of the tool, there are m and n stationary rectangular area sources of *variable* intensity distributed respectively over the top surface and the flank. Referring to Fig. 4(*a*), the steady temperature rise above ambient at the center of any subarea i' due to a rectangular area source situated at i (shown shaded) of uniform intensity $q_{t,i}$ is

$$\Delta\theta_{t,i'i} = \frac{q_{t,i}}{2\pi K_t} \int_{-w}^{3w} d\eta \int_{-\Delta x}^{\Delta x} \frac{d\xi}{\sqrt{(S_{ii'}-\xi)^2+\eta^2}} \quad \ldots [6]$$

which, upon integration, becomes

$$\frac{wq_{t,i}}{2\pi K_t}\left[\frac{|S_{ii'}+\Delta x|}{w}\left(\sinh^{-1}\frac{3w}{S_{ii'}+\Delta x}+\sinh^{-1}\frac{w}{S_{ii'}+\Delta x}\right)\right.$$
$$-\frac{|S_{ii'}-\Delta x|}{w}\left(\sinh^{-1}\frac{3w}{S_{ii'}-\Delta x}+\sinh^{-1}\frac{w}{S_{ii'}-\Delta x}\right)$$
$$+\sinh^{-1}\frac{S_{ii'}+\Delta x}{w}+3\sinh^{-1}\frac{S_{ii'}+\Delta x}{3w}$$
$$\left.-3\sinh^{-1}\frac{S_{ii'}-\Delta x}{3w}-\sinh^{-1}\frac{S_{ii'}-\Delta x}{w}\right] \quad \ldots\ldots [6a]$$

$S_{ii'} = 0$, when i' coincides with i. For a given $\dfrac{\Delta x}{w}$, the value of all the terms inside the bracket depends upon the ratio $\dfrac{S_{ii'}}{w}$, a dimensionless position parameter.

For convenience, one again writes

$$\Delta\theta_{t,i'i} = \frac{wq_{t,i}}{2\pi K_t}({}_iD_{i'}) \quad \ldots\ldots\ldots\ldots [6b]$$

where ${}_iD_{i'}$ is the sum of the terms inside the bracket of Equation

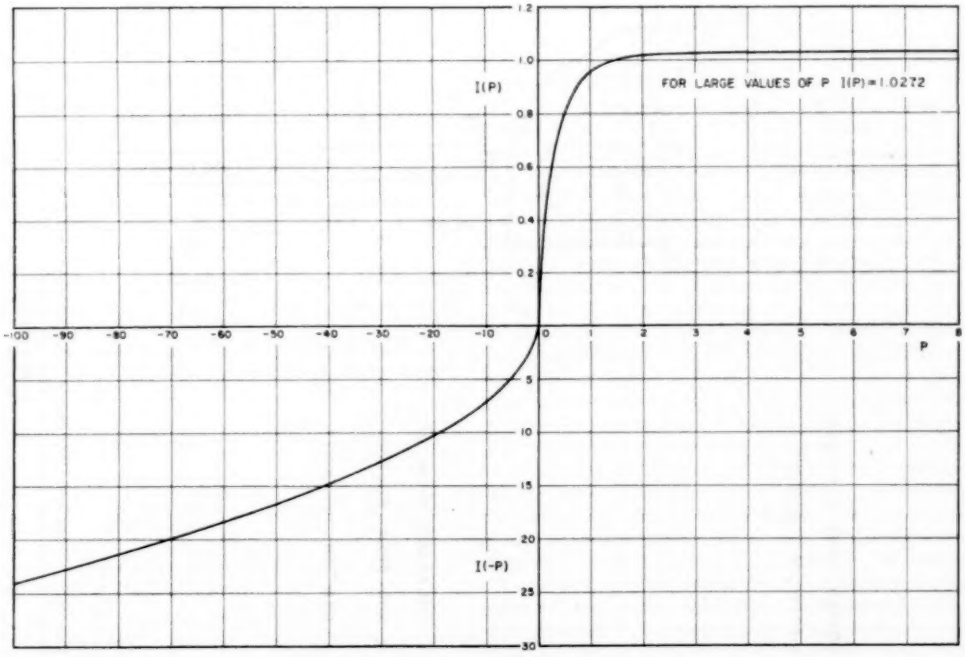


FIG. 3 VALUES OF THE DEFINITE INTEGRAL:

$$I(p)\int_0^p e^{-u}K_0(|u|)du, \quad I(-p)\int_0^p e^{u}K_0(|u|)du$$

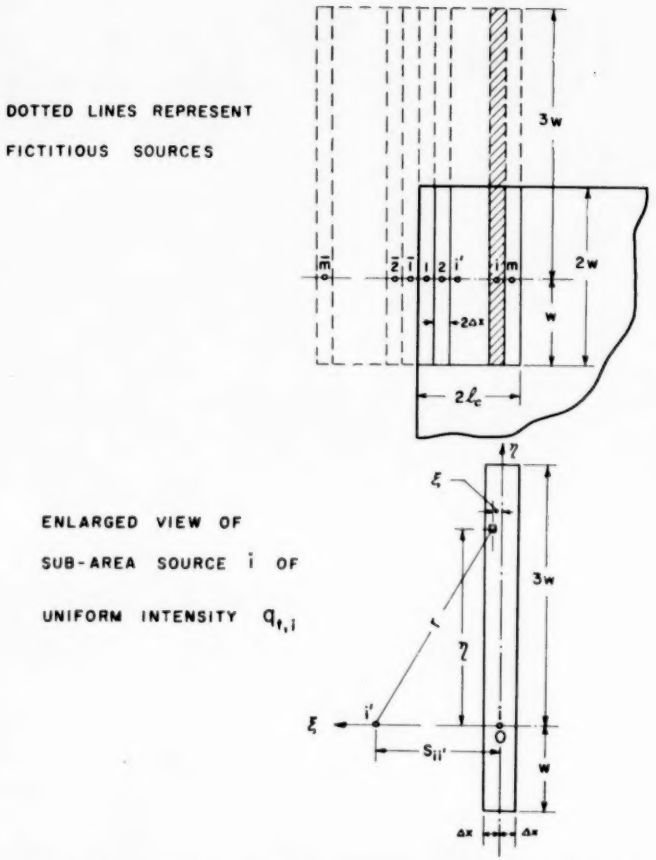


FIG. 4(a) REAL AND FICTITIOUS RECTANGULAR AREA SOURCES ON TOOL TOP SURFACE AND THE GEOMETRY USED IN THE COMPUTATION OF TEMPERATURE DISTRIBUTION

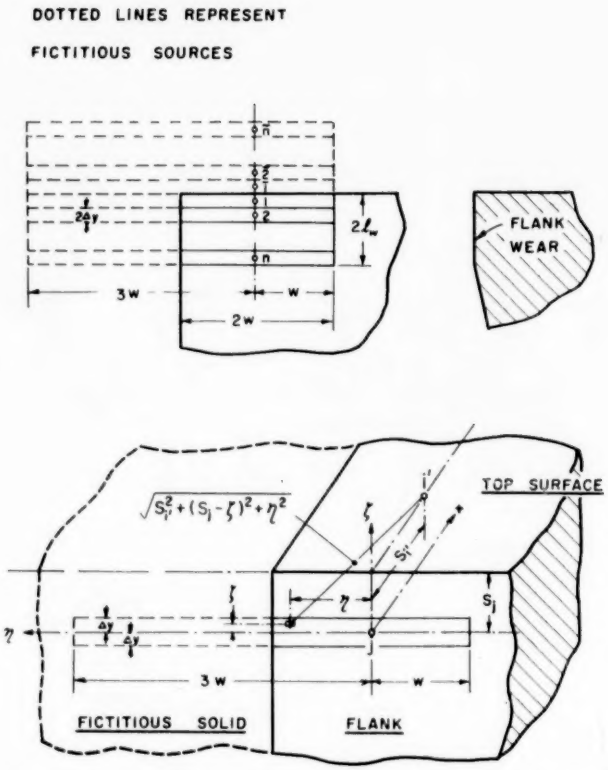


FIG. 4(b) REAL AND FICTITIOUS RECTANGULAR AREA SOURCES ON TOOL FLANK AND THE GEOMETRY USED IN THE COMPUTATION OF TEMPERATURE DISTRIBUTION

[6a]. The left subscript refers to the position of the source while the right subscript specifies the location the temperature of which is to be calculated. Clearly, i may vary from $\bar{m}, \ldots, \bar{1}, 1, \ldots m$. Position indexes with a bar refer to fictitious sources; $\bar{1}$ being the mirror image of 1, etc. i' varies from $1'$ to m'.

The temperature rise at i' due to an area source at the flank located at j may be formulated in a similar manner. Referring to Fig. 4(b), one writes

$$\Delta\theta_{f,i'j} = \frac{q_{f,j}}{2\pi K_t}\int_{-w}^{3w} d\eta \int_{-\Delta y}^{\Delta y} \frac{d\zeta}{\sqrt{[(S_{i'}^2 + (S_j - \zeta)^2 + \eta^2)]}} \quad \ldots\ldots [7]$$

It is not known to the authors that the double integral in Equation [7] can be expressed in the closed form. Since, in the present case, $0 < \frac{\sqrt{[S_{i'}^2 + (S_j - \zeta)^2]}}{w} < 1$, a close approximation can be given by

$$\begin{aligned}\Delta\theta_{f,i'j} = \frac{wq_{f,j}}{2\pi K_t}\Bigg\{ & \frac{y_2}{w}\ln\frac{12}{\left(\frac{y_2}{w}\right)^2 + \left(\frac{S_{i'}}{w}\right)^2} - \frac{y_1}{w}\ln\frac{12}{\left(\frac{y_1}{w}\right)^2 + \left(\frac{S_{i'}}{w}\right)^2} - 2\frac{S_{i'}}{w}\left(\tan^{-1}\frac{y_2}{S_{i'}} - \tan^{-1}\frac{y_1}{S_{i'}}\right) \\ & + \left(\frac{y_2}{w} - \frac{y_1}{w}\right)\left[2 + \frac{5}{18}\left(\frac{S_{i'}}{w}\right)^2 - \frac{41}{432}\left(\frac{S_{i'}}{w}\right)^4 + \frac{1825}{34{,}992}\left(\frac{S_{i'}}{w}\right)^6\right] + \left[\left(\frac{y_2}{w}\right)^2 - \left(\frac{y_1}{w}\right)^2\right] \\ & \left[\frac{5}{54} - \frac{41}{648}\left(\frac{S_{i'}}{w}\right)^2 + \frac{1825}{34{,}992}\left(\frac{S_{i'}}{w}\right)^4\right] - \left[\left(\frac{y_2}{w}\right)^5 - \left(\frac{y_1}{w}\right)^5\right]\left[\frac{41}{2160} - \frac{365}{11{,}664}\left(\frac{S_{i'}}{w}\right)^2\right] \\ & + \frac{1825}{244{,}944}\left[\left(\frac{y_2}{w}\right)^7 - \left(\frac{y_1}{w}\right)^7\right] + \ldots\Bigg\} \quad \ldots\ldots [7a]\end{aligned}$$

with $y_1 = S_j - \Delta y$ and $y_2 = S_j + \Delta y$.
For convenience, one writes

$$\Delta\theta_{f,i'j} = \frac{wq_{f,j}}{2\pi K_t}\,({}_jD_{i'}) \quad \ldots\ldots\ldots\ldots [7b]$$

where ${}_jD_{i'}$ is the sum of the terms inside the braces { } in Equation [7a]. j may vary from $\bar{n}, \ldots \bar{1}, 1, \ldots n$, i' from $1'$ to m' as before. By symmetry, ${}_jD_{i'} = {}_{\bar{j}}D_{i'}$.

The resultant temperature at location i' along OX (Fig. 2) of the tool-chip interface, due to heat sources both at the top surface and at the flank is thus

$$\theta_{t,i'} = \frac{w}{\pi K_t}\left\{\sum_{i=1}^{m} q_{t,i}\left(\frac{{}_iD_{i'} + {}_{\bar{i}}D_{i'}}{2}\right) + \sum_{j=1}^{n} q_{f,j}({}_jD_{i'})\right\} + \theta_0 \quad \ldots\ldots [8]$$

Following precisely the same procedure, the resultant temperature at j' along OY of the tool-work interface is

$$\theta_{f,j'} = \frac{w}{\pi K_t}\left\{\sum_{i=1}^{m} q_{t,i}({}_iD_{j'}) + \sum_{j=1}^{n} q_{f,j}\left(\frac{{}_jD_{j'} + {}_{\bar{j}}D_{j'}}{2}\right)\right\} + \theta_0 \quad \ldots\ldots [9]$$

Since, in this case, ${}_iD_{j'} = {}_{\bar{i}}D_{j'}$ by symmetry, ${}_iD_{j'}$ is the sum of the terms inside the braces of Equation [7a] with $S_{i'}$ replaced by $S_{j'}$ and y_1, y_2 by x_1 and x_2, respectively. $x_1 = S_i - \Delta x$, $x_2 = S_i + \Delta x$.

Equations [3] and [8] are merely two different expressions for the local interface temperature at i', hence $\theta_{c,i'} = \theta_{t,i'}$. Likewise, one concludes $\theta_{w,j'} = \theta_{f,j'}$, the former is given by Equation [5] and the latter by Equation [9]. One also observes that

$$q_{c,i} + q_{t,i} = q_c; \quad q_{w,j} + q_{f,j} = q_w \quad \ldots\ldots\ldots [10]$$

where q_c and q_w are, respectively, the rate of local heat liberation per unit area at the tool-chip and tool-work interfaces, calculable from tool force dynamometer measurements. Combining Equations [3], [5], [8], [9], [10] and rearranging gives

$$\sum_{i=1}^{m} q_{t,i}\left\{\frac{{}_iD_{i'} + {}_{\bar{i}}D_{i'}}{2} + 2\frac{K_t}{C_c\rho v_f w}({}_iI_{i'})\right\} + \sum_{j=1}^{n} q_{f,j}({}_jD_{i'}) = 2\frac{K_t}{C_c\rho v_f w} q_c \sum_{i=1}^{m} {}_iI_{i'} + \frac{\pi K_t}{w}(\bar{\theta}_s - \theta_0) \quad \ldots\ldots [11a]$$

and

$$\sum_{i=1}^{m} q_{t,i}({}_iD_{j'}) + \sum_{j=1}^{n} q_{f,j}\left\{\frac{{}_jD_{j'} + {}_{\bar{j}}D_{j'}}{2} + 2\frac{K_t}{C_w\rho v_c w}({}_jI_{j'})\right\} = 2\frac{K_t}{C_w\rho v_c w} q_w \sum_{j=1}^{n} {}_jI_{j'} \quad \ldots\ldots [11b]$$

Equations [11] constitute a set of $m + n$ simultaneous linear algebraic equations, the unknowns are m $q_{t,i}$'s and n $q_{f,j}$'s. With the division and distribution of heat flux at both interfaces evaluated, the local temperatures can be readily obtained from Equations [3] and [5] or [8] and [9].

Results and Discussion

Temperature distributions at the tool-chip and tool-work interface have been calculated for turning annealed AISI 4142 steel of 212 Bhn, using steel cutting-grade carbide at cutting speeds of 300, 496, and 700 fpm. A detailed numerical example has been given in reference (1) and need not be repeated here. Figs. 5 and 6 illustrate the influence on the two interface temperatures as flank wear develops. It is seen that, except for the initial drop, the calculated tool-chip interface temperature changes relatively little as flank wear increases, although a definite upward trend can be noticed. During cutting tests, the indicated temperature has been observed to be somewhat lower (~25 deg F) with a flank wear up to about 0.009 in. than with the initial sharp tool. At greater flank wear (the magnitude depends on the tool-work pair) the indicated temperature increases appreciably. This is not to be confused with the effect of flank wear on the calculated temperatures shown in Fig. 5. With flank wear the indicated temperature is *some* average of that at the many minute junctions at both tool-chip and tool-work contacts.

Unlike the tool-chip interface temperature, the temperature at the tool-work interface is greatly affected by the development of flank wear. This is shown in Fig. 6. While the absolute magnitude of the temperature shown may involve some error due to possible inaccuracies in the determination of frictional force at the tool flank, it is certain that (a) the tool flank temperature does not stay unchanged but increases appreciably as cutting proceeds, and (b) the distribution of temperature is nonuniform, with the maximum occurring at a location close to the point where the tool leaves contact with the workpiece. The nonuniformity becomes more pronounced as the flank wear gets larger.

The general trend of the development of flank wear with cut-

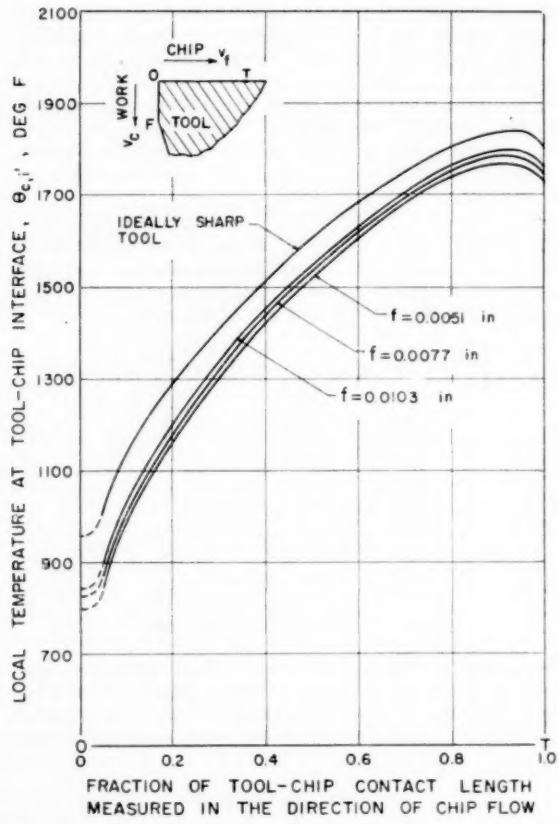


FIG. 5 EFFECT OF GROWTH OF FLANK WEAR ON TOOL-CHIP INTERFACE TEMPERATURE DISTRIBUTION

Work material: AISI 4142 steel, annealed, 212 Bhn
Tool material: Steel cutting-grade carbide
Tool shape: 0-6-7-7-10-0-0.015 in.
Cutting speed: V_s = 700 fpm
Feed: f_1 = 0.00632 ipr
Depth of cut: w_1 = 0.100 in.
Room temperature: θ_0 = 75 deg F

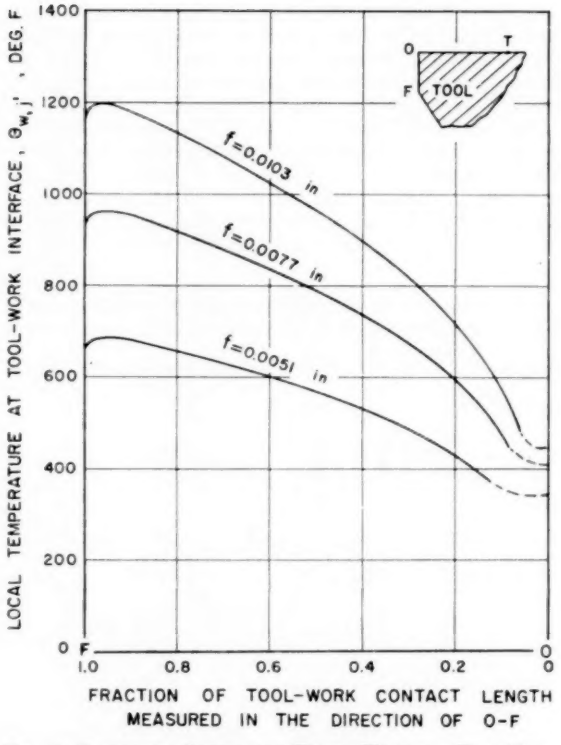


FIG. 6 EFFECT OF GROWTH OF FLANK WEAR ON TOOL-WORK INTERFACE TEMPERATURE DISTRIBUTION
(Cutting conditions same as in Fig. 5.)

ting time is shown in Fig. 7. Wear takes place at a rapid pace during the initial rubbing contact—known as the "break-in" wear. Local concentration of contact with the accompanying abnormally high contact stress is responsible for this phenomenon. From the study of the wear behavior of an SAE 1095 steel rider rubbing on a hardened steel disk, Burwell and Strang (10) reported that there was a sharp increase in wear rate when the contact stress exceeded a certain limiting value. The initial break-in wear of piston rings is attributed to this cause.

The portion AB of the wear curve is usually concave downward as shown. Frequently, it may be approximated by a straight line. Shaw and Dirke (11) have proposed a mechanism to explain this phenomenon. Beyond the point B, the wear rate rises rapidly. Although the flank temperature increases with wear land in the region AB, the wear rate is relatively insensitive to temperature change due to the low temperature level. In this region flank wear is predominantly of the abrasion type.

Shaw and Dirke's expression for adhesion wear has the form

$$W = K(nc)\frac{NL}{\sigma_y}$$

in which W is the wear volume when one surface slides past the other over a distance L, N the normal load, n the mean number of contact asperities in a unit length, c the mean height of a wear particle, K the probability that a contact will result in a wear particle, and σ_y the mean flow stress (or hardness) of a surface asperity. Following the consideration of inhomogeneities and imperfections of actual materials, Shaw and Dirke concluded, as a first approximation, that the product nc may be regarded as a constant. While both K and σ_y are generally temperature dependent, an increase in temperature will produce a significant increase in K and a decrease in σ_y *only* at certain high temperature levels according to the general theory of rate process (12). The rather abrupt increase in wear rate at and beyond the point B is thus explained.

In a recent study on the effect of edge conditions on flank wear development, it has been repeatedly observed that the presence of tool-work adhesion retards flank wear in a manner similar to the protection offered by a built-up edge to the top surface. As the flank wear increases beyond a certain limit, the magnitude of which depends on a particular tool-work pair, the flank built-up disappears, first in a region close to where the workpiece leaves contact with the tool. This observation supports the theoretical finding on temperature distribution depicted in Fig. 6.

Figs. 8 and 9 illustrate respectively the influence of cutting speed on temperature distribution over the tool-chip and tool-work interface, compared at a fixed flank wear of 0.0103 in. As is expected, both temperatures increase with cutting speed but the temperature at the flank increases more rapidly than that at the tool-chip interface. The latter fact can be readily understood The tool-chip interface temperature is composed of two parts; namely, the bulk temperature rise of the chip at the shear zone

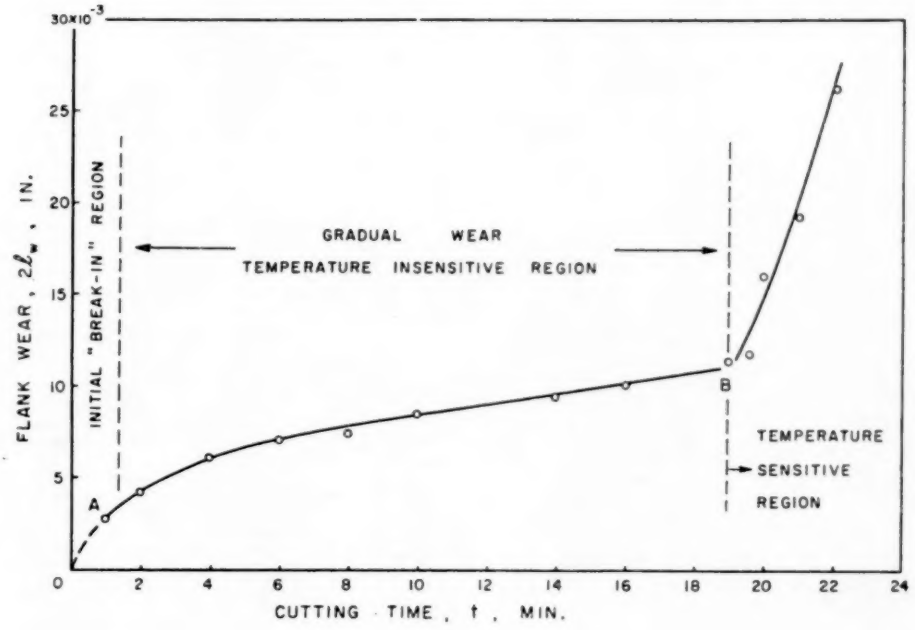


FIG. 7 GENERAL TREND OF FLANK WEAR DEVELOPMENT WITH CUTTING TIME INDICATING TEMPERATURE - SENSITIVE AND INSENSITIVE REGIONS

Work material: S-816 alloy
Tool material: steel cutting-grade carbide
Tool shape: 0-6-7-7-10-0-0.015 in.
Cutting speed: V_s = 50 fpm
Feed: f_1 = 0.00492 ipr
Depth of cut: w_1 = 0.100 in.)

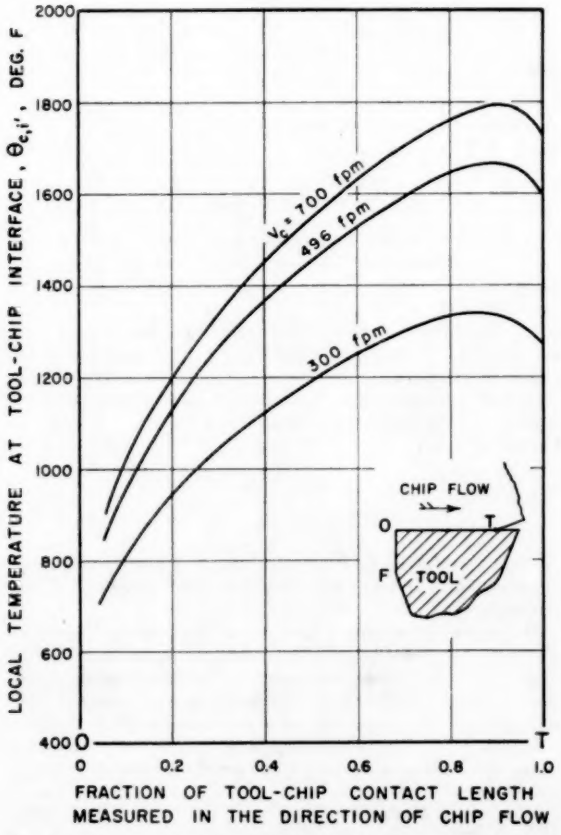


FIG. 8 TOOL-CHIP INTERFACE TEMPERATURE DISTRIBUTION FOR 3 CUTTING SPEEDS AT A FIXED FLANK WEAR OF 0.0103 IN.
(Other cutting conditions are same as in Fig. 5.)

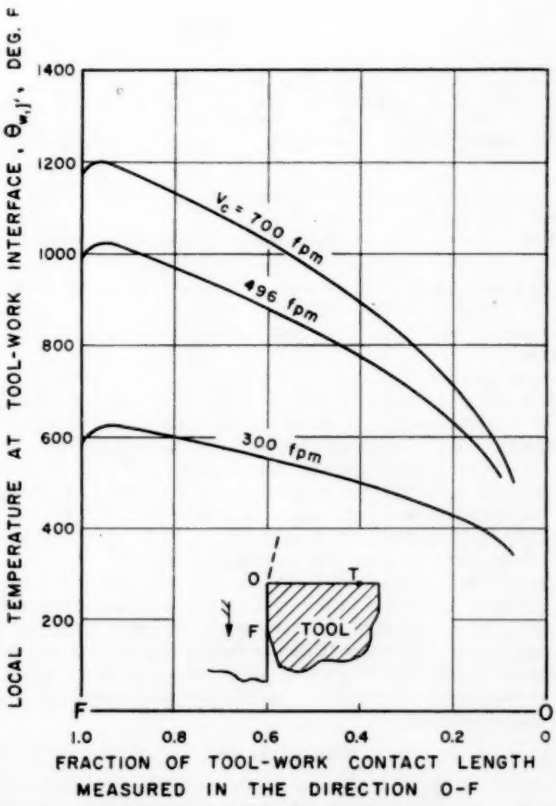


FIG. 9 TOOL-WORK INTERFACE TEMPERATURE DISTRIBUTION FOR 3 CUTTING SPEEDS AT A FIXED FLANK WEAR OF 0.0103 IN.
(Other cutting conditions are same as in Fig. 5.)

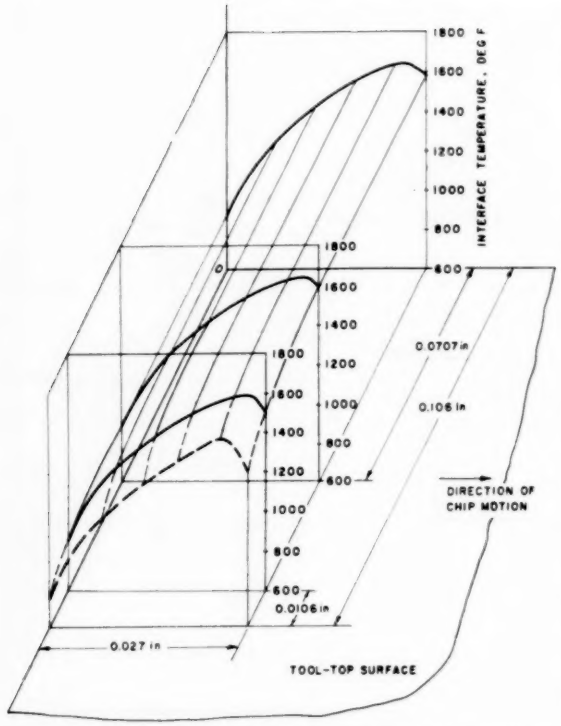


FIG. 10 THREE DIMENSIONAL TEMPERATURE DISTRIBUTION AT TOOL-CHIP INTERFACE
(Cutting conditions same as in Fig. 5, except the V_s = 496 fpm.)

and the additional temperature rise due to rubbing of the already heated chip on the top surface of the tool. When the speed of cutting is raised, the former is relatively unaffected (within the realm of type 2 chip formation) only the latter increases. On the other hand, the tool-work interface temperature rise is due solely to the frictional rubbing at the tool flank. Hence, an increase in cutting speed will result in a greater relative increase in temperature.

Variations in the heat-flux and temperature in a direction normal to chip flow were not considered in the preceding analysis. To evaluate the effect of such variations under the cutting conditions normally employed for sintered carbide tools the analysis was extended to a three-dimensional model. The method of approach was as already explained and a detailed computation was carried out for the case of an ideally sharp tool (2). The result is shown graphically in Fig. 10.

It is seen that under the conditions cited the temperature gradient in a direction normal to chip motion is insignificant except for a region close to the outer edge of the chip. However, with a great reduction in cutting speed, not only will the peak temperature shift toward the active cutting edge of the tool, but also the variation of temperature in the transverse direction will become greater. At the outer chip edge, the temperature may be reduced appreciably.

ACKNOWLEDGMENT

This work was sponsored by the Office of Ordnance Research, U. S. Army under Contract DA-11-022-ORD-1980. The authors express their appreciation to that Office for support of the program. Thanks are also due Dr. Y. H. Lee, General Electric Company, Schenectady, for his help in the computations: to Mr. D. L. Mykkanen, Department of Mechanical Engineering, University of Illinois, for valuable help in the research program; and to Miss Irene Cunningham for typing the manuscript.

BIBLIOGRAPHY

1 "Temperature and Heat Flux Distribution at Tool-Chip and Tool-Work Interface in Metal Machining," by B. T. Chao, K. J. Trigger, and Y. H. Lee, M. E. Technical Note: ORD-1121-1, University of Illinois, May, 1955.

2 "Three-Dimensional Temperature Distribution at the Tool-Chip Interface-Machining at High Speeds With Sintered Carbide Tools," by B. T. Chao and K. J. Trigger, M. E. Technical Report: ORD 1980-1, University of Illinois, March, 1956.

3 "Temperature Distribution at the Tool-Chip Interface in Metal Cutting," by B. T. Chao and K. J. Trigger, Trans ASME, vol. 77, 1955, pp. 1107–1121.

4 Discussion of the paper, Bibliography (3), by H. Blok.

5 "The Significance of the Thermal Number in Metal Machining," by B. T. Chao and K. J. Trigger, Trans. ASME, vol. 75, 1953, pp. 109–120.

6 "Shear-Plane Temperature Distribution in Orthogonal Cutting," by J. H. Weiner, Trans. ASME, vol. 77, 1955, pp. 1331–1341.

7 "On the Analysis of Cutting-Tool Temperatures," by E. G. Loewen and M. C. Shaw, Trans. ASME, vol. 76, 1954, pp. 217–231.

8 "Thermal Properties of Titanium Alloys and Selected Tool Materials," by E. G. Loewen, Trans. ASME, vol. 78, 1956, pp. 667–670.

9 "Moving Sources of Heat and the Temperature at Sliding Contacts," by J. C. Jaeger, Proceedings of the Royal Society, New South Wales, vol. 76, 1942, pp. 203–224.

10 "On the Empirical Laws of Adhesive Wear," by J. T. Burwell and C. D. Strang, *Journal of Applied Physics*, vol. 23, 1952, pp. 18–28.

11 "On the Wear of Cutting Tools," by M. C. Shaw and S. O. Dirke, paper presented at International Institution for Production Engineering Research, Milan, Italy, April, 1955. To be published in *Microtechnic*.

12 "The Theory of Rate Process," by S. Glasstone, K. J. Laidler, and H. Eyring, McGraw-Hill Book Company, Inc., New York, N. Y., 1941, Ch. 9.

Discussion

H. BLOK.[5] It is gratifying to note that, in following the suggestion of the writer in his discussion (reference 4) to the authors' previous paper (reference 3), a considerable degree of success has been achieved.

It would appear, however, that the authors missed a point in stating that "the definite integral which appears in Equation [1] cannot be expressed in closed form." In fact, it was shown previously by the writer[6] that the following two expressions, from which the authors' definite integrals $I(p)$ and $I(-p)$ can easily be found, hold good

$$\int e^{-u} K_0(|u|)\, du = ue^{-u} [K_0(|u|) - K_1(|u|)]$$

$$\int e^{u} K_0(|u|)\, du = ue^{u} [K_0(|u|) + K_1(|u|)]$$

In these expressions K_1 denotes the modified Bessel function of the second kind and first order (the integration constant has been omitted).

By means of the two expressions the relationship depicted in Fig. 3 can be verified. Thus, for instance, it can be found that for large values of p (approaching infinity) the definite integral does not approach 1.0272, as indicated in the figure, but unity. Admittedly, this inaccuracy is by no means appreciable. It is, therefore, expected that further verification will prove that the authors' ultimate results are accurate enough for the present purpose.

---

[5] Professor of Mechanical Engineering, University of Technology, Delft, Holland.

[6] See footnote 4 of "Dissipation of Frictional Heat" (in Dutch), by H. Blok, *Voordrachten, Koninklijk Institut van Ingenieurs*, vol. 2, 1950, pp. 84–104.

As diagrams depicting the division and distribution of the frictional heat in the contact areas could be very instructive for future calculations by others, it is suggested that the authors in their reply give at least one such diagram.

E. G. Loewen.[7] The writer is intrigued with the idea the authors have presented concerning the noniterative temperature calculation. While the temperature distribution shown in Fig. 5 has the expected general shape, one wonders if the position of the temperature peak, so very near the heel of the chip, still agrees with what the iterative procedure, reference (3) of the Bibliography of the paper, would have predicted. In other words, is this new method fully equivalent to the old one?

Until we can manage to collect some good experimental data on the important problem of shear-plane temperature distribution, it is difficult to quarrel with the authors' assumption that it is uniform. I think that this is one of the most important gaps in our knowledge of the problem.

For some years now the writer has been wondering about the amount of energy dissipated in friction between the work and the clearance face of the tool. He has tried, with no success at all, to figure out how the amounts of heat generation on tool-chip and tool-work interfaces can be measured separately, with force-dynamometer data. He is most disappointed to note that the authors give no indication at all of how they managed to do this. If they did it simply by taking differences observed between sharp and worn tools, then it is doubted whether the result is too significant. Perhaps they can now shed some light on this matter.

[7] The Taft-Peirce Manufacturing Company, Woonsocket, R. I. Assoc. Mem. ASME.

Authors' Closure

The authors wish to thank the discussers for their interest in the paper. They are particularly indebted to Professor Blok for pointing out the possibility of expressing the definite integral in Equation [1] in terms of known tabulated functions. This not only results in a better accuracy but also facilitates the computation by eliminating the tedious work involved in graphical integration. The reference cited by Professor Blok (footnote 6) is in Dutch and, unfortunately, had escaped the authors' attention.

Several diagrams depicting the division and distribution of frictional heat over both areas of contact can be found in the Bibliography (1). Two of such diagrams are reproduced in Figs. 11 and 12 of this closure for further reference. In Fig. 11, $q_{t,i}/q_c$ represents the fraction of local tool-chip interface heat transferred to the tool. For an ideally sharp tool, heat is actually flowing from the tool into the chip over a small distance in the proximity of the cutting edge. This is in agreement with the result reported by Rapier.[8] Nevertheless, an ideally sharp tool does not exist in practice because of the extremely rapid "break-in" wear. As flank wear develops, there is an initial radical change in the direction of heat flow near the cutting edge as shown in Fig. 11. Further development of flank wear results

[8] Discussion by A. C. Rapier of "Some Factors Affecting Wear on Cemented Carbide Tools," by E. M. Trent, Proceedings of the Institution of Mechanical Engineers, London, England, vol. 166, 1952, pp. 64–74.

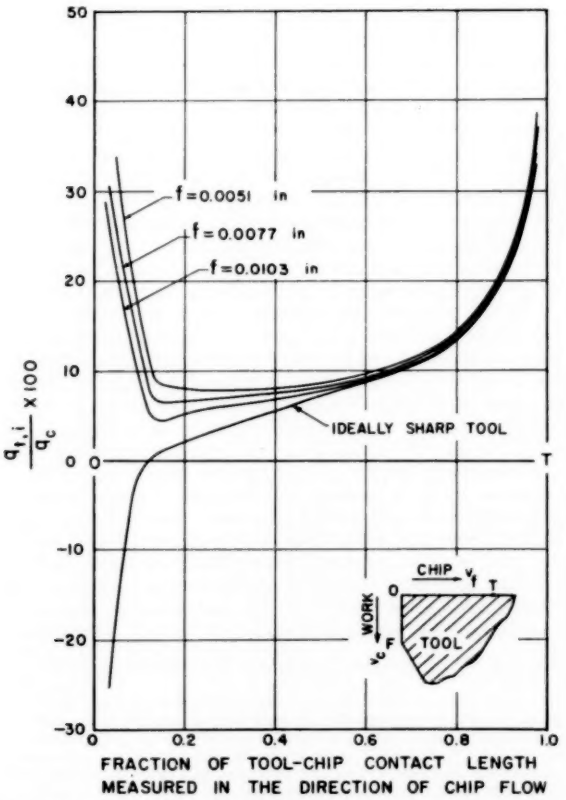


Fig. 11 Effect of Growth of Flank Wear on Tool-Chip Heat-Flux Distribution
(Cutting conditions same as in Fig. 5.)

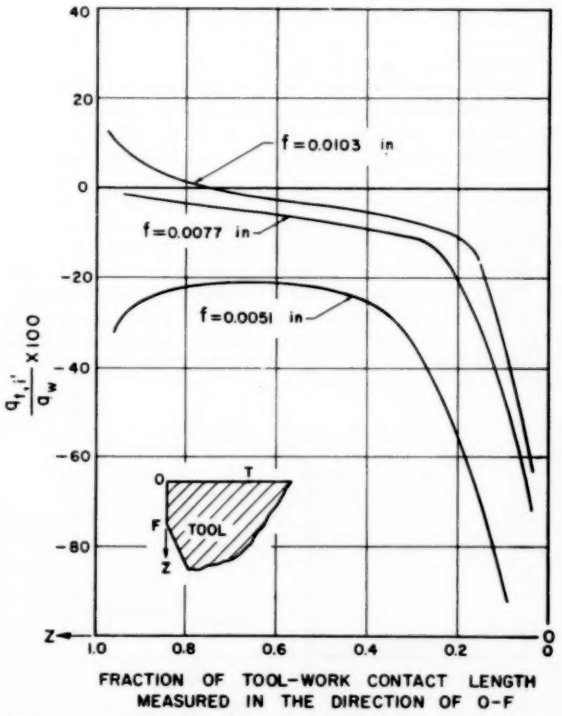


Fig. 12 Tool-Work Interface Heat-Flux Distribution at Three Flank Wear Measurements
(Cutting conditions same as in Fig. 5.)

in a slight decrease of the proportion of tool-chip interface heat which flows into the tool.

Fig. 12 illustrates the distribution of heat flux at the tool-work interface for three flank-wear measurements. It is seen that with the exception of the largest flank wear considered where $f = 0.0103$ in. and in a region close to the point F where the tool leaves contact with the workpiece, all the frictional heat is transferred to the workpiece. This is conceivable since the workpiece has a much lower bulk temperature and effectively serves as a heat sink.

Dr. Loewen has raised a question concerning the equivalency of the method of calculation presented in this paper and that of the original one given in the Bibliography (3). This has been explained at the beginning of the paper under Introduction. All the major assumptions used in the current analysis have also been listed and explained. Some of them were adopted simply because of lack of information on the particular subject. The authors have never claimed that the shear-plane temperature *is* uniform. This problem has been studied by Hahn,[9] Weiner,[10] Vieregge,[11] and others including the authors.[12] Perhaps the analysis of Weiner is, theoretically, the best. Unfortunately, Weiner's expression for shear-plane energy distribution lacks agreement with the experimental data of Vieregge. It is under this circumstance that the assumption of a uniform chip-bulk temperature is made. While the result will certainly be dependent upon the reliability of such an assumption, the procedure of computation presented here will not be affected. Oliver Heaviside once said: "Shall I refuse my dinner because I do not fully understand the process of digestion?" However, the authors do agree with Dr. Loewen that there is much need for experimental data on the subject of shear-plane temperature distribution.

[9] "On the Temperature Developed at the Shear Plane in the Metal-Cutting Process," by R. S. Hahn, Proceedings of the First U. S. National Congress of Applied Mechanics, 1951, pp. 661–666.

[10] See Bibliography (6).

[11] "Energieverteilung und Temperatur bei der Zerspannung," by G. Vieregge, *Werkstatt und Betrieb*, vol. 86, 1953, pp. 691–703.

[12] See Bibliography (5).

Transient Interface Temperatures in Plain Peripheral Milling

By D. E. McFERON[1] and B. T. CHAO,[2] URBANA, ILL.

The analytical calculation of tool-chip interface temperature has been extended to the plain peripheral milling process. The solution of the equations presented enables an investigation of the effects of material and process variables to be made from fundamental cutting data. It was found that the intermittent nature of this cutting process increases the percentage of heat flow from the tool-chip interface into the tool as compared with single-point turning. A slight increase in the workpiece temperature has a more pronounced influence on the tool-chip interface temperature than would a similar increase in the tool temperature.

Introduction

THE removal of unwanted metal in the form of chips is the objective of most metal-cutting operations. In forming these chips under the usual cutting conditions, the tools are subjected to high local stresses and temperatures. Much information has been obtained during the past half century concerning forces and temperatures together with their relationships to other machining variables, particularly for the case of single-point turning (1).[3] Not nearly as much information is available for the other machining processes. In 1949 it was reported (2) that less was known about the cutting of metals by milling than by any other machining process. A survey of the current literature shows that this condition still prevails, primarily because of the transient nature of the process which makes experimental determinations relatively difficult.

As early as 1925 experimental methods of determining the average temperature between the tool and the chip were devised (3, 4, 5). In 1938 Schallbroch, Schaumann, and Wallichs (6) developed empirical expressions relating tool life to the temperature developed in machining. Trigger and Chao (7) recently have given a more detailed explanation of this effect. An analytical method of relating tool forces and other cutting variables to the temperature was developed in 1951 (8). Since that time a few refinements to the procedure have been made (9) and experimental evidence has accumulated to verify its use. Again, however, most of this information has been gathered for single-point tools.

In order to study the application of these procedures to a transient cutting condition the process of plain peripheral milling was chosen. Plain peripheral milling is the process in which the intermittent cutting action between the rotating cutter and the stationary workpiece takes place on the outer periphery of the cutter and the thickness of the chip formed increases to a maximum during the cut. Thus, not only is the cutting action intermittent, but also the cutting conditions are changing continuously. Determination of the transient tool-chip interface temperature, both analytically and experimentally, is the primary objective of this paper.

The Tool-Chip Interface Temperature Equation

Model. Considering an ideally sharp tool, i.e., a tool which has not developed a "wear land" on the clearance surface, there are two sources of heat in the formation of a continuous chip during orthogonal cutting. These heat sources consist of the shear plane (OA) and the tool-chip interface (OB) as shown in Fig. 1. The temperature rise at the interface is caused by plastic deformation during the formation of the chip at the shear plane and the frictional rubbing of the chip on the top surface of the tool as it subsequently passes off. Procedures have been evolved for the calculation of these temperatures under steady-state conditions existing in single-point turning (8, 9). In plain peripheral milling, however, the conditions of cutting are changing continuously during chip formation and the process is transient in nature.

As has been pointed out by Martellotti (10) the path of a tooth in plain milling is characterized by a trochoidal curve. However, for the cutting conditions typically recommended (11) for milling steel with carbide cutters, this path can be approximated closely by an arc of a circle. This simplification results in a less complicated expression for the computation of the instantaneous "uncut" chip thickness, t_1. For the conditions encountered in this investigation, a further simplification, that the "uncut" chip thickness varied linearly with angular displacement of the cutter, resulted in less than $^1/_2$ per cent error in the determination of this quantity; hence the model adopted was that of a long triangular chip having a height-to-base ratio of approximately 1:1000.

Shear Plane Temperature. The shear plane can be represented as an oblique band source of heat moving in the workpiece. Previous analysis (9) has shown that for the case of calculating the average temperature, which is of interest here, a good approximation can be obtained by replacing this oblique source by a plain slider moving at the shear velocity in the surface of the workpiece which is regarded as semi-infinite in extent. The average temperature rise above the initial uniform temperature

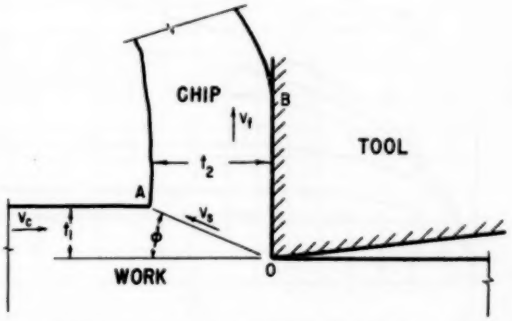


Fig. 1 Geometry of Chip Formation and Heat Sources in Orthogonal Cutting; Type 2 Chip and Ideally Sharp Tool

[1] Associate Professor of Mechanical Engineering, University of Illinois.

[2] Professor of Mechanical Engineering, University of Illinois.

[3] Numbers in parentheses refer to Bibliography at end of paper.

Contributed by the Research Committee on Metal Processing and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 6, 1956. Paper No. 56—A-89.

of a semi-infinite solid over the area of contact of the moving band source at any time T after the beginning of the movement has been given by Jaeger (12) as

$$\Delta\bar{\theta} = \frac{q\kappa}{2\sqrt{(\pi)}KLV}\int_0^{\frac{V^2T}{2\kappa}}\left\{\Phi\left[\frac{2L+u}{\sqrt{(2u)}}\right]+\Phi\left[\frac{|2L-u|}{\sqrt{(2u)}}\right]-2\Phi\left[\left(\frac{u}{2}\right)^{1/2}\right]\right\}du \quad [1]$$

In this expression, q is the rate of heat generation per unit area, $L = (Vl)/(2\kappa)$, V is the velocity of the source, l is one half the source length in the direction of sliding, κ is the thermal diffusivity of the material ($= K/(c\rho)$, where K is the thermal conductivity, and $c\rho$ the volumetric specific heat), and Φ is the integral of the error function

$$\Phi(x) = \int_0^x \operatorname{erf}\beta\, d\beta = x\operatorname{erf}x - \frac{1}{\sqrt{(\pi)}}(1-e^{-x^2})$$

In order to investigate the rate at which steady-state temperature is approached, a dimensionless plot of Equation [1] was made and is shown in Fig. 2. This figure gives the time T necessary to approach the steady-state values of temperature for the parameter, $0.2 \leqslant L \leqslant 5.0$.

As a typical example for the cutting conditions encountered in the experimental phase, for a cutting speed of 264 sfpm

$$L_s\left(=\frac{v_s l_s}{2\kappa}\right)$$

is 0.34 after a cutter rotation of $\psi = 1$ deg, and the corresponding value of $(V^2T)/(2\kappa)$ at this time is 148. From Fig. 2 it is seen that for the given value of L steady temperature is reached for all practical purposes at $(V^2T)/(2\kappa) \sim 2.5$. Hence, for the speeds and feeds used in the present investigation, the transient shear-plane temperature may be calculated to a good approximation by using simple expressions derived for steady-state conditions. It is recognized that the foregoing analysis does not hold rigorously for the milling process in that the contact length varies continuously with time. However, an examination of Fig. 2 shows that the time required to reach equilibrium decreases for smaller values of L.

In considering the temperature developed at the shear plane, Equation [1] must be modified with respect to q, the heat flux available to cause the temperature rise. The product $F_s v_s$ gives the gross rate of energy liberation on the shear plane in which

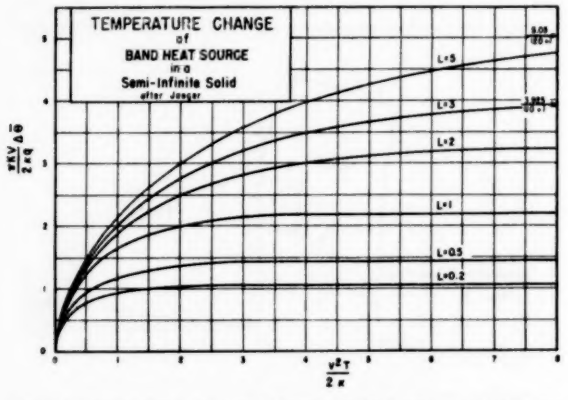


FIG. 2 TRANSIENT AVERAGE TEMPERATURE RISE OF A MOVING-BAND HEAT SOURCE IN THE SURFACE OF A SEMI-INFINITE SOLID

F_s is the shearing force and v_s the shearing velocity. For the deformations encountered in machining, almost all of the energy at the shear plane is liberated as sensible heat with only a small percentage used in permanent lattice deformation (13); however, the heat liberated is divided between the slider (chip) and the stationary workpiece. A procedure outlined by Blok (14) for the determination of λ_s, the fraction of the shear-zone heat going into the workpiece, is used in the subsequent analysis.

Since the values of L will be between 0.2 and 5.0 for our cutting conditions, Fig. 3, for the average steady temperature over a moving band source within these limits of the parameter L will be used. The temperature rise at the shear plane considered from the point of view of the workpiece is

$$\Delta\bar{\theta}_s = \frac{2n\kappa\lambda_s q_s}{\pi K v_s} \quad [2]$$

where n is the ordinate in Fig. 3.

From the standpoint of the chip, the temperature rise is due to an amount of heat $(1-\lambda_s)\, q_s A_s$ being released on the shear plane per unit time. The amount of material traversing the shear plane per unit time is most conveniently given by $t_1 w_1 v_c \rho$. The thickness and width of the "uncut" chip are, respectively, t_1 and w_1; v_c is the cutting speed and ρ the density of the work material. It is evident then that the temperature rise on the shear plane can also be given as

$$\Delta\bar{\theta}_s = \frac{(1-\lambda_s)\, q_s A_s}{t_1 w_1 v_c c\rho} \quad [3]$$

Since the shear-plane area $A_s = t_1 w_1 \csc\phi$, Equation [3] may be written as

$$\Delta\bar{\theta}_s = \frac{(1-\lambda_s)\, q_s \csc\phi}{v_c c\rho} \quad [3a]$$

Equations [2] and [3a] can be equated to give an expression for λ_s which, on simplifying, becomes

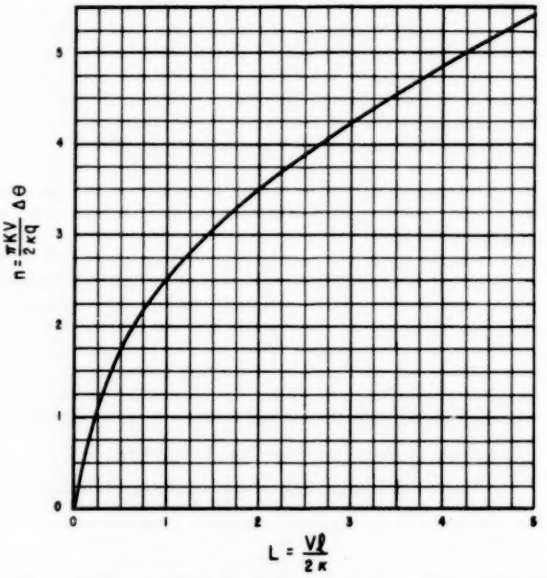


FIG. 3 STEADY AVERAGE TEMPERATURE RISE OF A MOVING-BAND HEAT SOURCE IN THE SURFACE OF A SEMI-INFINITE SOLID—AFTER JAEGER

$$\lambda_s = \frac{1}{1 + \frac{2n}{\pi} \frac{\cos\ (\phi - \alpha)}{\cos\ \alpha} \sin\ \phi} \quad \ldots\ldots\ldots\ldots [4]$$

By making the tool rake angle α equal to zero this expression can be simplified further to

$$\lambda_s = \frac{1}{1 + \frac{2n}{\pi} \sin\ \phi \cos \phi} \quad \ldots\ldots\ldots\ldots\ldots [4a]$$

Once λ_s has been determined from Equation [4a], substitution into either Equation [2] or [3a] will give the average temperature rise at the shear zone.

Tool-Chip Interface Temperature. In evaluating the temperature rise at the tool-chip interface in milling due to frictional rubbing, the division of heat can again be calculated from two points of view. For the tool, one has a stationary source of varying intensity acting over a variable area of contact. The chip, however, "sees" the heat source as a moving-plane slider. Again by equating the average temperature obtained from these two viewpoints it is possible to calculate the division of heat between the two elements and then the average interface temperature.

In order to calculate the temperature rise on the tool surface due to friction at the tool-chip interface, use will be made of Kelvin's integration of the Fourier heat-conduction equation as found in the standard textbooks on heat conduction (15). This equation gives the temperature rise at any point (x, y, z) in an infinite conducting solid initially at a uniform temperature, t units of time after the instantaneous liberation of a finite quantity of heat Q, from a point source located at the origin. It is given as

$$\Delta\theta = \frac{Q\kappa}{8K\ (\pi\kappa t)^{3/2}}\ e^{-\frac{x^2+y^2+z^2}{4\kappa t}} \quad \ldots\ldots\ldots\ldots [5]$$

For a semi-infinite solid with no surface heat loss and restricting our attention to the surface ($z = 0$), Kelvin's equation becomes

$$\Delta\theta = \frac{Q\kappa}{4K\ (\pi\kappa t)^{3/2}}\ e^{-\frac{x^2+y^2}{4\kappa t}} \quad \ldots\ldots\ldots\ldots [5a]$$

That the orthogonal tool can be considered as a quadrant of an infinite solid in so far as temperature calculations are concerned has been shown in previous investigations (8, 9). It also has been shown that the heat loss at the clearance surface can be neglected for the usual cutting conditions. Consequently, the temperature rise at any point (x, y) in the top surface of the tool due to a finite quantity of heat liberated instantaneously and uniformly over an area $l_c \times 2w$ as shown in Fig. 4 is identical to that in a semi-infinite solid with the additional contribution of a fictitious source of the same intensity located at the mirror image of the actual source as is illustrated by the dotted lines in the same figure. Hence, this temperature rise may be obtained from Equation [5a] by summing up the effects of the individual point sources having strengths

$$\frac{Q}{2l_c \times 2w}\ dx'\ dy'$$

over the entire area $2l_c \times 2w$ as

$$\Delta\theta = \frac{1}{4K_t\ (\pi\kappa_t\ t)^{3/2}} \frac{Q\kappa_t}{2l_c \times 2w} \int_{-l_c}^{+l_c} dx' \int_{-w}^{+w} e^{-\frac{(x-x')^2+(y-y')^2}{4\kappa_t t}}\ dy' \ldots [6]^4$$

where x' and y' are the co-ordinates of the point heat source.

When integrated, Equation [6] becomes

$$\Delta\theta = \frac{1}{4K_t} \sqrt{\left(\frac{\kappa_t}{\pi t}\right)} \frac{Q}{2l_c \times 2w} \left\{\text{erf}\ \frac{x + l_c}{2\sqrt{(\kappa_t\ t)}} - \text{erf}\ \frac{x - l_c}{2\sqrt{(\kappa_t\ t)}}\right\} \left\{\text{erf}\ \frac{y + w}{2\sqrt{(\kappa_t\ t)}} - \text{erf}\ \frac{y - w}{2\sqrt{(\kappa_t\ t)}}\right\} \ldots [6a]$$

where $$\text{erf}\ x = \frac{2}{\sqrt{(\pi)}} \int_0^x e^{-u^2}\ du$$

To find the average temperature existing over an area $2l_{c.t} \times 2\ w$, t units of time after the heat is released, Equation [6a] will be integrated over the desired area and divided by the same area. Thus

$$\Delta\bar{\theta} = \frac{1}{2l_{c.t} \times 2w} \int_{-l_{c,t}}^{+l_{c,t}} dx \int_{-w}^{+w} \Delta\theta\ dy \ldots\ldots [7]$$

Since the temperature distribution is symmetrical with respect to the X and Y-axes, Equation [7] can be modified slightly to read

$$\Delta\bar{\theta} = \frac{1}{l_{c.t} \times w} \int_0^{l_{c.t}} dx \int_0^w \Delta\theta\ dy \ldots\ldots\ldots [7a]$$

Substituting the value of $\Delta\theta$ from Equation [6a] and simplifying results in

$$\Delta\bar{\theta} = \frac{1}{(c\rho)_t} \sqrt{\left(\frac{\kappa_t t}{\pi}\right)} \frac{Q}{2l_c \times 2w} \frac{1}{l_{c.t} \times w} \left\{\Phi\left[\frac{l_{c.t} + l_c}{2\sqrt{(\kappa_t\ t)}}\right] - \Phi\left[\frac{l_{c.t} - l_c}{2\sqrt{(\kappa_t\ t)}}\right]\right\} \Phi\left[\frac{2w}{2\sqrt{(\kappa_t\ t)}}\right] \ldots [8]$$

This expression gives the average temperature existing over an area $2l_{c.t} \times 2\ w$, t units of time after an amount of heat Q is instantaneously and uniformly released over an area $2l_c \times 2w$.

In the actual milling process, the interface heat is liberated continuously and at a variable rate. The average temperature at t units of time after contact begins is given by

$$\Delta\bar{\theta} = \frac{1}{(c\rho)t} \sqrt{\left(\frac{\kappa_t}{\pi}\right)} \frac{1}{l_{c.t} \times w} \int_0^t q\sqrt{(t - \tau)} \left\{\Phi\left[\frac{l_{c.t} + l_{c.\tau}}{2\sqrt{[\kappa_t\ (t - \tau)]}}\right] - \Phi\left[\frac{l_{c.t} - l_{c.\tau}}{2\sqrt{[\kappa_t\ (t - \tau)]}}\right]\right\} \Phi\left[\frac{2w}{2\sqrt{[\kappa_t\ (t - \tau)]}}\right] d\tau \ldots [9]$$

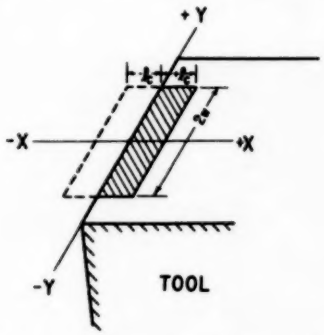


FIG. 4 REAL AND FICTITIOUS SOURCE OF HEAT ON TOP SURFACE OF ORTHOGONAL TOOL

[4] The subscript t refers to the tool

where $l_{c,t}$ = length of contact at time t
$l_{c,\tau}$ = instantaneous length of contact, $0 \leqslant \tau \leqslant t$
q = instantaneous heat flux flowing into tool over contact area $l_{c,\tau} \times 2w$

With Equation [9] it is possible to calculate an average interface temperature rise considering the amount of heat being released and the tool-chip contact area to be functions of time as they are in the peripheral milling process. In performing the integration of Equation [9], the contact time is subdivided into a finite number of intervals. Details of this procedure are given in the sample calculation.

The heat which is released at the tool-chip interface is shared by the chip and the tool. Again, Blok's (14) partition principle will be used to determine λ_i, which is the fraction of the interface energy going into the tool. The use of this method is permissible for reasons previously given.

The interface heat acts as a band source moving in the surface of the chip. From Fig. 3, a value of n can be obtained as a function of

$$L_c \left(L_c = \frac{v_f l_c}{2\kappa} \right)$$

which has been found to be greater than 0.2 and less than 5.0 for the cutting conditions selected. An expression for the average interface temperature from the point of view of the chip can be formulated as

$$\bar{\theta}_i = \frac{2n(1-\lambda_i)\, q_i}{\pi (c\rho)_c v_f} + \bar{\theta}_s \quad \ldots\ldots\ldots\ldots\ldots [10]$$

where $\bar{\theta}_s = \Delta\bar{\theta}_s + \theta_0$, θ_0 is the ambient temperature.

The average interface temperature, calculated from the point of view of the tool according to Equation [9] becomes

$$\bar{\theta}_i = \frac{1}{(c\rho)_t} \sqrt{\left(\frac{\kappa_t}{\pi}\right)} \frac{1}{l_{c,t} \times w} \int_0^t \lambda_i q_i \sqrt{(t-\tau)} \left\{ \left[\Phi \frac{l_{c,t} + l_{c,\tau}}{2\sqrt{[\kappa_t (t-\tau)]}} - \Phi \frac{l_{c,t} - l_{c,\tau}}{2\sqrt{[\kappa_t (t-\tau)]}} \right] \right\} \Phi \frac{2w}{2\sqrt{[\kappa_t (t-\tau)]}}\, d\tau + \theta_0 \ldots [11]$$

A numerical method was used to evaluate Equation [11] in terms of λ_i which was then determined by equating [10] and [11]. Once λ_i has been determined, substitution into either of these expressions will give the value of $\bar{\theta}_i$.

It is pertinent that the foregoing procedure involves an approximation as the average temperatures are superimposed. Strictly, such procedure can be applied only to point values.

Experimental Equipment

The experimental program was carried out on a heavy-duty, horizontal, plain milling machine. This machine, which is rigidly built for milling with carbide cutters has a 20-hp main-drive motor and a separate 3-hp motor for the feed mechanism.

The forces of milling were measured by means of a three-component strain-gage dynamometer. In this design the strain elements were in the form of octagonal rings as suggested by Cook, Loewen, and Shaw (16). The strain gages were connected into three independent circuits in the form of complete bridges of eight gages to each bridge. The bridge output was connected to a d-c wide-frequency-response preamplifier and thence to a cathode-ray oscilloscope. The cathode-ray oscilloscope (CRO) was equipped with a P-7 long-persistence screen and bezel illumination which greatly facilitated the photography of the transient forces in milling.

The average tool-chip interface temperature was measured by means of the well-known tool-work thermocouple technique (3, 4, 5, 17). A special milling cutter as shown in Fig. 5 was constructed to provide the connections required for this arrangement. The connection from the tool tip was brought to the preamplifier through an electrically insulated metal wheel rotating in a mercury bath. No difficulty was experienced with the mercury bath since the wheel speed was relatively low (about 100 fpm maximum). Several checks of the contact resistance gave readings of less than 1 ohm. The workpiece for the cutting-temperature tests was electrically insulated and clamped rigidly in a vise mounted on the milling machine. A shielded wire connected the workpiece to the preamplifier. The general experimental arrangement is shown in Fig. 6.

The tool material used in these tests was a tungsten, titanium, tantalum cemented carbide of Kennametal grade K3H which was brazed to tool holders of AISI 4140 steel. The carbide blanks used were approximately $^5/_{16}$ in. $\times$ $^5/_8$ in. $\times$ 1 in. in size. For the work material AISI 4140 steel, quenched and tempered to 270 Brinell hardness number was used. It was chosen predominantly because the Metal Cutting Laboratory had extensive lathe-cutting data for similar steels with which to compare the milling results. Inasmuch as all of the tests were of short duration, the total used up only a small amount of metal from the

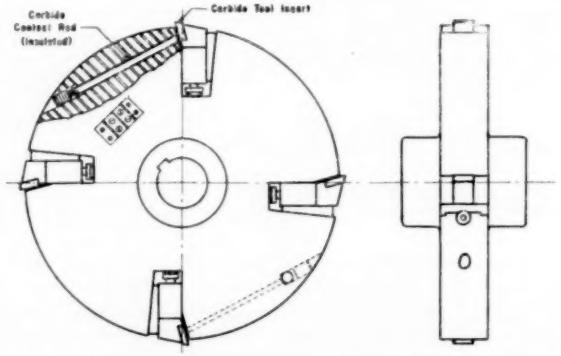


Fig. 5 Special Peripheral Milling Cutter for Temperature Measurement

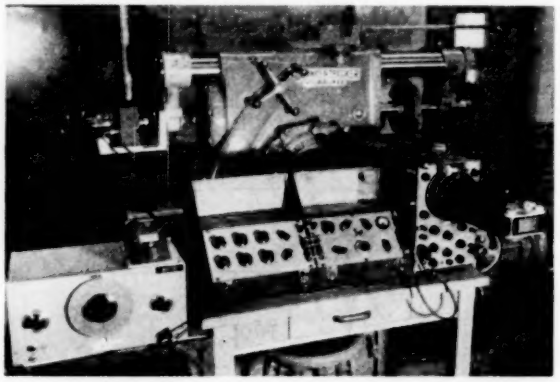


Fig. 6 Experimental Arrangement for Measuring Transient Forces and Tool-Chip Interface Temperature in Milling

test bar thus tending to minimize any metallurgical variations. An average test run used up about 1 in. in length of the workpiece.

Test Procedure

In selecting the cutting conditions commercial recommendations (11) for this tool-work combination were followed in so far as it was practicable. The feed per tooth was selected to be approximately 0.005 in. with a 0.050 in. depth of cut. Secondary cutting conditions of 0.003 in. feed per tooth of 0.100 in. depth of cut were chosen so as to form the same "uncut" chip thickness (10) as the previous combination. Cutting speeds ranged from about 180 sfpm below which the chip became discontinuous to about 400 sfpm above which the tool edge failed rapidly. As is customary in commercial milling with carbide cutters, all of the tests were made dry; i.e., cut in air. During cutting the feed was always in a direction counter to the cutter rotation. This results in what is commonly known as "up milling" or "conventional milling."

The tools were ground with 0-deg axial and radial rake angles to simplify the force measurements and computations and a 6-deg peripheral relief angle. A workpiece width of $^1/_2$ in. was chosen so that the $^5/_8$-in-wide tool would overhang the cut on each side to provide orthogonal cutting conditions. In this manner only the end cutting edge was active during the cut.

In order to prevent overlapping of the instantaneous force and temperature indications on the CRO, it was set on driven sweep and triggered by a rotating mechanical switch attached to the milling-machine spindle. This switch had an adjustable contact position so that the CRO sweep could be initiated at a fixed geometric spacing in relationship to the actual cut. A series switch kept the rotating switch from triggering the sweep when not desired. The oscilloscope trace was recorded on 35-mm film and enlarged for subsequent analysis.

During the test runs no record was taken until the cut was well established and chips of the desired geometry were being produced. These chips were collected for measurement of the actual thickness which was necessary in order to compute the chip-thickness ratio r_t. They also served the useful purpose during testing of being quite sensitive indicators of any damage to the cutting edge. If any such damage was suspected the test was stopped immediately. In measuring the final chip thickness, representative chips were mounted by recasting them in drilled holes of precast lucite metallurgical mounting cylinders. This allowed a section normal to the direction of cutting to be polished and measured. A typical section is shown in Fig. 7.

Extensive lathe-cutting data (18) have shown that the chip-thickness ratio for given materials cut at constant speed varies with the feed. The variation has been shown to be a straight line on logarithmic co-ordinates. To obtain the instantaneous

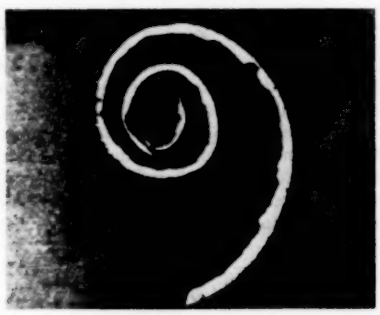


Fig. 7 Typical Cross Section of a Milling Chip

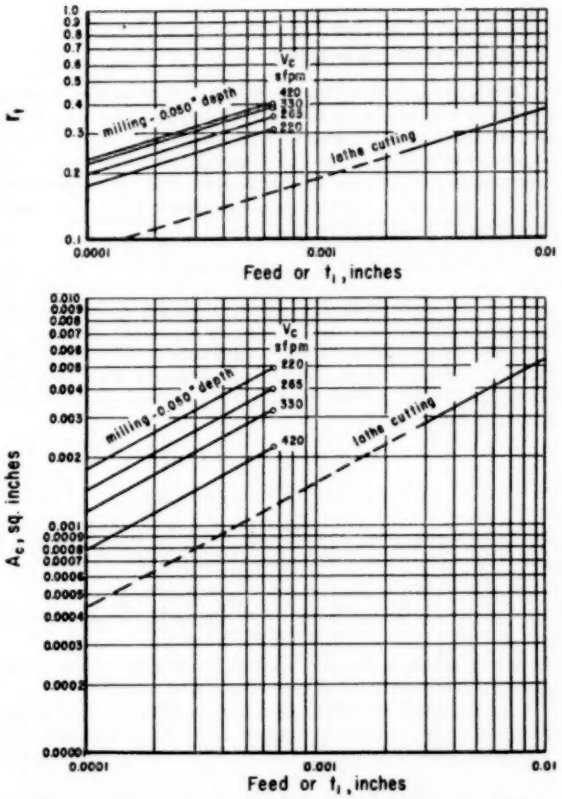


Fig. 8 Graphs Used in Approximate Determination of Instantaneous Chip-Thickness Ratio and Apparent Tool-Chip Contact Area

chip-thickness ratio during a cut, the r_t at maximum "uncut" chip thickness was obtained from direct measurement and plotted on logarithmic co-ordinates along with lathe cutting data. Since all of the slopes for the lathe-cutting data were approximately the same, a line was drawn through the experimentally determined point for milling and with this slope allowing intermediate points to be read where required. This procedure is illustrated in Fig. 8 and in the sample calculation.

The maximum area of contact between the tool top surface and the chip was obtained by photographing the etched surface after a test run. This procedure has been described previously (8) for single-point turning. Since this gives only the maximum area of contact a procedure identical to that given in the foregoing was used for determining contact areas at less than the maximum position.

Sample Calculation

The calculation procedure and a comparison with the experimental results can best be made by carrying out a sample calculation. Values obtained from a typical set of cutting conditions are as follows:

Tool: K3H carbide, 0 deg radial rake angle, 0 deg axial rake angle, 6 deg peripheral relief angle
Work: AISI 4140, quenched and tempered, 270 Bhn
Cutting speed: $V_c = 330$ fpm, $v_s = 3960$ ipm
Diameter of cutter: $D = 10.30$ in.

Depth of cut: $d = 0.050$ in.
Width of cut: $w = 0.503$ in.
Angle of contact: $\psi = 8$ deg

Typical oscillograms of the forces obtained under these cutting conditions are given in Fig. 9. In this figure the forces as recorded are for the horizontal and vertical components of the forces while it was desired to use the tangential (F_c) and radial (F_t) components for purposes of calculation. The relationship between these various components can be seen in Fig. 10 and from trigonometry it can be shown that

$$F_c = F_h \cos \psi - F_v \sin \psi \quad [12]$$

$$F_t = F_v \cos \psi + F_h \sin \psi \quad [13]$$

F_c and F_t are the tool force components in Merchant's (19) analysis. Other values calculable from his work, which are somewhat simplified from the general expressions since the rake angle α is 0 deg, are

$$\text{Shear angle } \phi = \tan^{-1} r_t \quad [14]$$

$$\text{Shearing strain } \epsilon = \tan \phi + \cot \phi \quad [15]$$

The instantaneous "uncut" chip thickness can be obtained from Equation [20] in Martellotti's (10) paper as

$$t_1 = f_t \frac{(2Rd - d^2)^{1/2} - f_t}{[R^2 - 2f_t(2Rd - d^2)^{1/2} + f_t^2]^{1/2}} \quad [16]$$

In this equation, f_t is the feed per tooth, R is the radius of the cutter, and d is the depth of cut, all measured in inches.

For the cutting conditions used in these tests the chip can be assumed to be triangular, varying uniformly from zero to a maximum value as given by Equation [16]. This approximation results in less than 4 per cent error at the beginning of the cut and decreases to a negligible amount at $\psi > 4$ deg.

TABLE 1 CHIP THICKNESS RATIO AND CONTACT AREA

ψ, deg	r_t	A_c, sq in.
1	0.22	0.001
2	0.27	0.0015
3	0.30	0.0018
4	0.32	0.0022
5	0.34	0.0025
6	0.36	0.0027
7	0.37	0.0030

TABLE 2

ψ deg	T $\times 10^5$ min	t_1 $\times 10^3$ in.	Fc, lb	Ft, lb	ϕ	s	$2l_s$ $\times 10^3$ in.	l_c $\times 10^3$ in.
1	2.27	0.78	62.4	61.3	12°25′	4.76	3.63	2.0
2	4.54	1.55	96.0	78.3	14°54′	4.02	5.66	2.9
3	6.81	2.32	122.2	99.6	16°29′	3.68	8.17	3.7
4	9.08	3.10	142.6	122.7	17°45′	3.44	10.17	4.3
5	11.35	3.88	166.3	141.9	18°53′	3.27	12.00	4.9
6	13.62	4.65	187.2	162.7	19°36′	3.16	13.88	5.4
7	15.89	5.43	199.6	186.3	20°18′	3.07	15.63	5.9

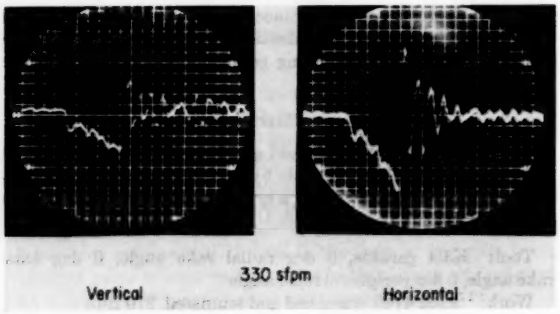


FIG. 9 TYPICAL OSCILLOSCOPE RECORDS OF TRANSIENT TOOL FORCES DURING MILLING

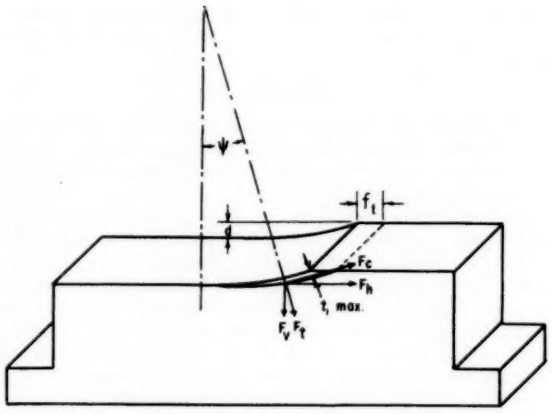


FIG. 10 MILLING GEOMETRY

The values of the chip-thickness ratio and the contact area at $\psi = 8$ deg were measured experimentally, as previously described, and intermediate values obtained from Fig. 8. These values are shown in Table 1.

An examination of Fig. 1 will show that the length of the shear zone is

$$2l_s = t_1 \csc \phi \quad [17]$$

The length of contact on the tool top surface is given by

$$l_c = A_c/w_1 \quad [18]$$

The time required for the cutter to rotate through an angle ψ is given as

$$T = \frac{D \times \pi}{360} \times \frac{\psi}{t_c} \quad [19]$$

Values calculated from these relations are shown tabulated in Table 2 for the example under consideration.

Average Shear-Plane Temperature. It is first necessary to calculate a value of the parameter

$$L_s = \frac{v_s l_s}{2\kappa_w}$$

in which

$$v_s = \frac{v_c}{\cos \phi} \quad (\alpha = 0 \text{ deg})$$

From Fig. 3 a value of n corresponding to the calculated L_s can be found and, when substituted into Equation [4*a*], gives the partition fraction, λ_s. This fraction represents the proportion of the shear-zone heat which is transmitted to the workpiece.

Once this fraction has been determined, the average shear-plane temperature rise can be obtained from either Equation [2] or [3*a*]. A convenient expression can be derived by substituting equivalent relationships into Equation [3*a*]. This results in

$$\Delta\bar{\theta}_s = \frac{(1 - \lambda_s) S_s}{J c\rho \sin \phi \cos \phi} \quad [20]$$

in which S_s is the dynamic flow stress of the work material.

For the example being considered, these calculated values are shown in Table 3.

Tool-Chip Interface Temperature. The average temperature of the tool-chip interface will be determined at 1-deg intervals within the angle of contact. The instantaneous values of the length of contact, $l_{c,\tau}$ will be obtained by dividing the instantaneous area of contact A_c, as given in Table 1 by the width of cut w_1. The rate of heat generation per unit area is

$$q_i = \frac{F v_f}{J l_{c,\tau}\, w_1} \qquad [21]$$

in which the friction force, $F = F_t$ since $\alpha = 0$ deg, chip-flow velocity, $v_f = v_c r_b$, and J is the mechanical equivalent of heat.

The thermal properties of the tool material are[5]

$$K_t = 0.0228 \frac{\text{Btu}}{\text{min} - \text{sq in.} - \frac{\text{deg F}}{\text{in.}}}$$

$$\rho_t = 0.401 \frac{\text{lb}_\text{M}}{\text{in.}^3}$$

$$c_t = 0.066 \frac{\text{Btu}}{\text{lb}_\text{M} - \text{deg F}}$$

Since for the tool top surface the area of contact was continuously varying as was the intensity of the heat released, and the expression for the instantaneous average temperature contained rather involved functions Equation [11], resort was made to numerical means for evaluating the temperature at intervals of the cutter rotation. In this manner the cumulative effect of the heat input over the varying area of contact could be determined. For the interval from $\psi = 0$ deg to $\psi = 1$ deg this numerical procedure gave a value of the integral part of Equation [11] of 10.82 Btu $\text{min}^{1/2}$/sq in. for the example being considered. Substituting the other known values into Equation [11] gives

$$\Delta\bar{\theta}_{i(1,1)} = 429\lambda_{i(0\text{-}1)}$$

or

$$\theta_{i(1,1)} = 429\,\lambda_{i(0\text{-}1)} + 75 \text{ F}$$

(ambient temperature is 75 F) where $\Delta\bar{\theta}_{i(1,1)}$ is the average interface temperature rise at $\psi = 1$ deg resulting from heat being released from $\psi = 0$ deg to $\psi = 1$ deg.

The interface temperature also can be evaluated from Equation [10]. After calculating

$$L_c = \frac{v_f l_c}{2\kappa_c} = 0.407, \quad n = 1.6$$

The specific heat of the chip material c_c is determined from Fig. 11 at the tool-chip interface temperature. This involves an estimation of the temperature and possibly a correction after the first calculation if the original estimate proves to be in error by more than 25 F. Differences of less than 25 F have a negligible effect on the calculated interface temperatures. Estimating $\bar{\theta}_i$ to be 650 F, $c_c = 0.139$ Btu/lb_M deg F. Substituting other values in Equation [10] results in

$$\Delta\,\bar{\theta}_{i(1,1)} = 169.7\,[1 - \lambda_{i(0\text{-}1)}]$$

and

$$\bar{\theta}_{i(1,1)} = \Delta\,\bar{\theta}_{i(1,1)} + \bar{\theta}_s = 169.7[1 - \lambda_{i(0\text{-}1)}] + 681 \text{ deg F}$$

Equating the two expressions for $\bar{\theta}_{i(1,1)}$ gives

$$\lambda_{i(0\text{-}1)} = 1.297$$

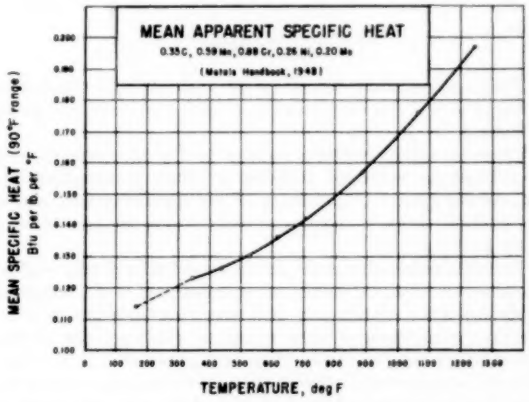


FIG. 11 VARIATION OF SPECIFIC HEAT OF AN ALLOY STEEL WITH TEMPERATURE

[5] The values of thermal conductivity K_t and density ρ_t are taken from reference (20). The specific heat c_t is taken from reference (21).

TABLE 3 AVERAGE SHEAR-PLANE TEMPERATURE CALCULATION

ψ, deg	L_s	n	λ_s	$(1 - \lambda_s)$	F_s, lb	S_s, psi $\times 10^{-3}$	c^*, Btu/lb_M-deg F	$\Delta\bar{\theta}_s$, deg F	$\bar{\theta}_s$, deg F
1	0.368	1.41	0.842	0.158	47.9	264	0.124	606	681
2	0.580	1.82	0.776	0.224	72.7	256	0.126	694	769
3	0.844	2.30	0.716	0.284	88.9	217	0.126	681	756
4	1.058	2.60	0.676	0.324	98.4	194	0.125	655	730
5	1.255	2.80	0.647	0.353	111.6	186	0.125	649	724
6	1.460	3.02	0.622	0.374	122.1	176	0.124	640	715
7	1.650	3.20	0.601	0.399	123.0	159	0.123	602	677

c^* is evaluated from Fig. 11 at a temperature midway between the initial and final shear plane temperature.
ρ is substantially independent of deformation and temperature change for the range encountered in cutting.

TABLE 4 CALCULATED TOOL-CHIP INTERFACE TEMPERATURES IN MILLING, DEG F

Heat source	ψ 1°	2°	3°	4°	5°	6°	7°	λ_i
0–1	556	68	39	29	20	17	14	1.297
1–2		637	114	67	48	36	31	0.882
2–3			612	123	73	53	41	0.657
3–4				576	106	76	54	0.510
4–5					579	137	83	0.447
5–6						551	136	0.373
6–7							524	0.317
$\Delta\,\bar{\theta}_i$	556	705	765	795	826	870	883	
$\bar{\theta}_i$	631	780	840	870	901	945	958	

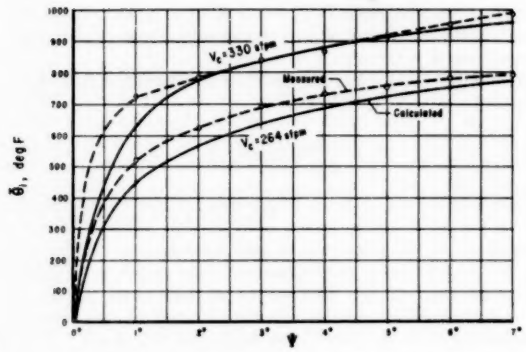


FIG. 12 COMPARISON OF CALCULATED AND MEASURED TRANSIENT TOOL-CHIP INTERFACE TEMPERATURES IN MILLING

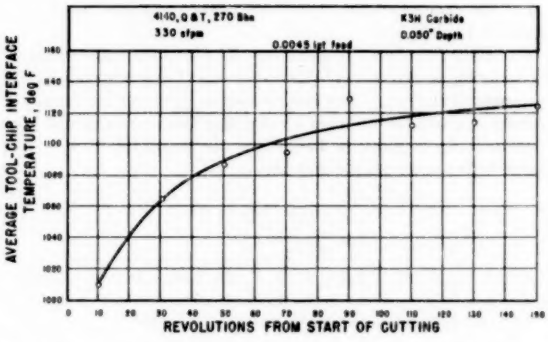


FIG. 13 MEASURED CHANGE IN MAXIMUM TOOL-CHIP INTERFACE TEMPERATURE DURING MILLING

and thus

$$\bar{\theta}_{i(1,1)} = 631 \text{ F}$$

The effect of the heat released during the period from $\psi = 0$ deg to $\psi = 1$ deg on later times can be evaluated in a similar manner by substituting appropriate values of t for the upper limit in Equation [11]. Carrying out this process for $\psi = 2$ deg results in

$$\bar{\theta}_{i(2,1)} = 68 \text{ F}$$

These and other calculations for the angle of contact from $\psi = 0$ deg to $\psi = 7$ deg for the example being considered are shown in Table 4.

These calculated results are shown in Fig. 12 compared with experimentally measured tool-chip interface temperatures for two cutting speeds. Considering all the approximations used in the analysis, some of them admittedly crude, the agreement is fairly good.

From the calculation it should be noted that not only is all of the interface frictional energy being transferred to the tool at the beginning of the cut but some heat from the chip is also ($\lambda_i > 1$). This is the antithesis of lathe cutting in which a major portion (85 to 95 per cent) of the interface heat passes off with the chip. The situation occurs in milling because of contact between the hot chip and the cool tool at the beginning of the cut. The value of λ_i decreases rapidly as the cut proceeds, but does not approach the value obtained in the steady-state lathe-cutting process.

SUMMARY AND CONCLUSIONS

A procedure and the necessary equations have been developed for the computation of the average, transient tool-chip interface temperature in plain peripheral milling. The analysis makes possible a prediction of the effect of changing cutting conditions and material properties on the tool-chip interface temperature. This temperature is particularly important with respect to tool wear. Since milling is a fairly complex operation to analyze, many of the results of variations in the process are not intuitively predictable.

Based on calculations, a uniform increase in the tool-tip temperature of 100 F would result in less than a 10 deg rise in the average tool-chip interface temperature; however, a corresponding increase in the workpiece temperature would result in a tool-chip interface temperature increase of almost 90 F. The sample calculations presented were based on both the tool tip and the workpiece being at ambient temperature. This would tend to give temperatures lower than actual after the first cut. The change in the measured maximum interface temperature on successive cuts is shown in Fig. 13.

Data on the thermal properties of the tool material are quite meager. A recent study (22) on the thermal conductivities of several carbide tool materials at elevated temperature showed that a composition similar to the cutting tool used in these tests had little variation in thermal conductivity with temperature. No information was found concerning the specific heat at other than room temperature. A calculation made to ascertain the effect on the interface temperature if the specific heat increased with temperature showed that a 100 per cent increase in specific heat resulted in less than 4 per cent decrease in the calculated interface temperature. This increase in specific heat would increase the amount of interface heat flowing into the tool by 15 to 20 per cent. This should cause the temperature on successive cuts, as shown in Fig. 13, to approach an asymptotic value more rapidly.

ACKNOWLEDGMENTS

The authors express their appreciation for the continual support and interest of Prof. K. J. Trigger during the experimental phases of the study and for his valuable contributions in reviewing of the original thesis. Thanks are due to Kennametal, Inc., Latrobe, Pa., for the carbide tool materials used in the investigation and to Miss Irene Cunningham for typing the final manuscript.

This paper is based on a part of a doctoral dissertation submitted to the Graduate College of the University of Illinois, February, 1956.

BIBLIOGRAPHY

1 "Manual on Cutting of Metals," ASME, New York, N. Y., 1952.

2 "Carbide Cutting Tools," by W. Baker and J. Kozacka, American Technical Society, Chicago, Ill., 1949, p. 297.

3 "Thermoelectric Measurement of Cutting Tool Temperatures," by H. Shore, *Journal of the Washington Academy of Sciences*, vol. 15, March 4, 1925, pp. 85–88.

4 "The Measurement of Cutting Temperatures in the Turning of Ingot Iron," by K. Gottwein, *Maschinenbau*, vol. 4, 1925, pp. 1129–1135.

5 "The Measurement of Cutting Temperatures," by E. G. Herbert, Proceedings of The Institution of Mechanical Engineers, London, England, vol. 1, 1926, pp. 289–329.

6 "Testing for Machinability by Measuring Cutting Temperatures and Tool Wear," by H. Schallbroch, H. Schaumann, and R. Wallichs, Vortrage der Hauptversammlung 1938 der deutschen Gesellschaft für Metallkunde, VDI-Verlag, 1938, pp. 34–38.

7 "The Mechanism of Crater Wear of Cemented Carbide Tools," by K. J. Trigger and B. T. Chao, ASME Paper No. 55—SA-11.

8 "An Analytical Evaluation of Metal-Cutting Temperatures," by K. J. Trigger and B. T. Chao, Trans. ASME, vol. 73, 1951, pp. 57–68.

9 "On the Analysis of Cutting-Tool Temperatures," by E. G. Loewen and M. C. Shaw, Trans. ASME, vol. 76, 1954, pp. 217–231.

10 "An Analysis of the Milling Process," by M. Martellotti, Trans. ASME, vol. 63, 1941, pp. 677–700.

11 "CSM Cutters, Catalogue C-11A," Kearney and Trecker Corporation, Milwaukee, Wis., 1946.

12 "Moving Sources of Heat and the Temperature at Sliding Contacts," by J. C. Jaeger, Proceedings of the Royal Society of New South Wales, vol. 76, 1942, pp. 203–224.

13 "The Latent Energy Remaining in a Metal After Cold Working," by G. I. Taylor and H. Quinney, Proceedings of the Royal Society of London, England, series A, vol. 143, 1934, pp. 307–326.

14 "Theoretical Study of Temperature Rise at Surfaces of Actual Contact Under Oiliness Lubricating Conditions," by H. Blok, Proceedings of the General Discussion on Lubrication and Lubricants, The Institution of Mechanical Engineers, London, England, 1938, pp. 222–235.

15 "Introduction to the Mathematical Theory of the Conduction of Heat in Solids," by H. S. Carslaw, Macmillan and Company, London, England, 1921, p. 150.

16 "Machine Tool Dynamometers," by N. H. Cook, E. G. Loewen, and M. C. Shaw, *American Machinist*, vol. 98, 1954, pp. 125–129.

17 "Progress Report No. 1 on Tool-Chip Interface Temperatures," by K. J. Trigger, Trans. ASME, vol. 70, 1948, pp. 91–98.

18 "Thermophysical Aspects of Metal Cutting," by B. T. Chao and K. J. Trigger, Trans. ASME, vol. 74, 1952, pp. 1039–1054.

19 "Basic Mechanics of the Metal-Cutting Process," by M. E. Merchant, Trans. ASME, vol. 66, 1944, p. A-168.

20 "Kennametal Cemented Carbide Products," Catalog 54, Kennametal, Inc., Latrobe, Pa., 1953, p. 62.

21 "Efficient Milling," Metal Cutting Institute, New York, N. Y., 1950, p. 18.

22 "Thermal Properties of Titanium Alloys and Selected Tool Materials," by E. G. Loewen, Trans. ASME, vol. 78, 1956, pp. 667–670.

Experience With Chromium-Molybdenum-Vanadium Steel in High-Temperature Bolting Applications

By R. G. MATTERS[1] and C. D. DICKINSON[2]

Chromium-molybdenum-vanadium steels, ASTM A 196, grade B 14 or slight modifications of this material have been used widely for high-temperature bolting since about 1938. In the temperature range 850 to 1000 F these materials have significantly higher relaxation resistance than have the other ferritic steels covered by ASTM A 196. Service performance of chromium-molybdenum-vanadium bolts in turbines operating with inlet steam temperatures below 1000 F has been generally satisfactory. However, within the past two years bolts have failed in two central station turbines. This paper covers the examination of the failed bolts, the analysis of the cause of failure, and recommendations to avoid failures of this type.

Bolting Failure, Turbine A

A 107,000-kw turbine, designated turbine A, operating on 1000 F steam had been in service for approximately 4 years. In early operation of the unit the two stop-valve cover joints leaked. This was a simple bolted joint with relatively short bolts tapped into the stop-valve body. The bolts were tightened to a strain of 0.0018 in. per in. Occasionally, the operating temperature of the unit varied fairly rapidly over a range of about 100 deg F. It was believed that the differential thermal expansion during the temperature swings overstressed the short bolts and caused the joint to leak.

Approximately two years prior to failure, the bolting for this joint was redesigned for increased elasticity in order to minimize the effect of temperature variation. As shown in Fig. 1, the effective elasticity of the bolts was increased by spacers installed above the valve cover. The bolts were tightened to the usual 0.0018 in. per in. strain, or 41,500 psi stress at temperature. After the new bolts and spacers had been in service for two years and as part of a general inspection of the unit, the joint was broken and the bolts were examined by ultrasonic inspection methods but were not removed from the valve body for magnaflux inspection. The ultrasonic tests indicated that none of the bolts contained serious cracks. The bolts were retightened to 0.0018 in. per in. strain and the unit was put back on the line. After two months of satisfactory service, a steam leak occurred in one of the stop-valve cover joints, and upon examination 8 out of 16 bolts in this stop valve and 4 out of 16 bolts in the other stop valve were found to have failed. At the time of failure these bolts had been in service for approximately 17,500 hr. At the time the failures occurred other bolts in the unit were examined for cracks by the magnaflux method. No cracks were found in any bolts which did not have spacers to increase the effective elasticity of the bolts.

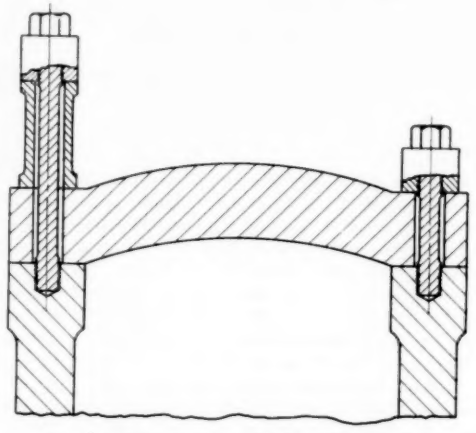


Fig. 1 Schematic diagram of bolted joint without and with spacers

The bolts failed with no visible ductility at the root of the first thread engaged in the valve body. The fractured surfaces of three of the bolts are shown in Fig. 2. Dark oxide films around the periphery of the fractures indicated that cracks had started from the roots of the threads some time prior to final failure of the bolts.

The chemical analyses of the bolts met the requirements of ASTM A 196, grade B 14, and indicated that all of the bolts were probably from a single heat of material with the following chemical composition:

	Per cent		Per cent
Carbon	0.46	Molybdenum	0.37
Manganese	0.52	Silicon	0.28
Chromium	1.03	Vanadium	0.25

Diameter measurements along the lengths of three bolts indicated that plastic deformation occurred primarily near the fixed end of the bolt as shown in Fig. 3. The major portion of the bolt shanks had undergone little if any plastic deformation. Hardness traverses along the lengths of three bolts indicated a distinct drop in hardness from the free to the fixed ends. The results of the diameter and hardness measurements strongly indicated that a rather steep temperature gradient existed along the bolt in service.

An estimate of the temperature distribution along the bolts was made based on the observed hardness values and the tempering behavior of the material. Previous tempering studies (1)[3]

[1] Assistant Director of Research, Research Laboratories, Allis-Chalmers Manufacturing Company, Milwaukee, Wis.

[2] Research Metallurgist, Research Laboratories, Allis-Chalmers Manufacturing Company, Milwaukee, Wis. Present Address: Mallory-Sharon Titanium Corporation, Niles, Ohio.

Contributed by the Joint ASME-ASTM Committee on the Effect of Temperature on Properties of Metals and presented at the Annual Meeting, New York, N. Y., November 25–30, 1956, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, November 1, 1956. This paper was not preprinted.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

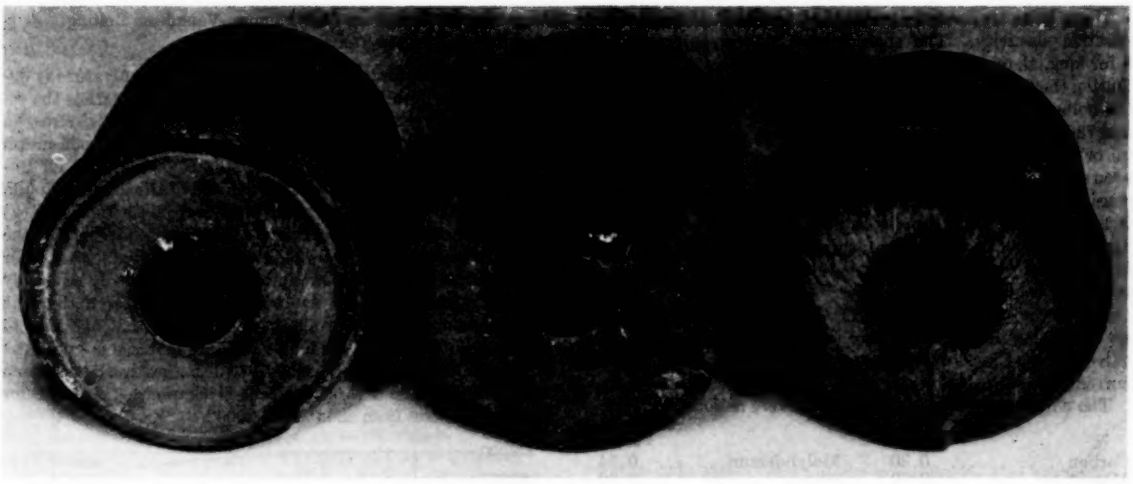


Fig. 2 Fractures of failed bolts from Turbine A

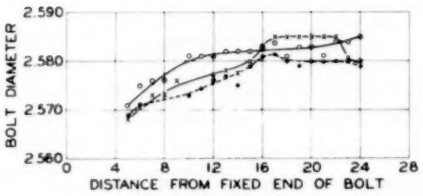


Fig. 3 Variation of diameter along bolt shank—bolts from Turbine A

have shown that the effect of time and temperature on hardness after tempering can be correlated by a parameter

$$P = (20 + \log t)T \quad \ldots\ldots\ldots\ldots\ldots\ldots [1]$$

where P is a temperature-time parameter which is correlated with hardness, t is time in hours, and T is absolute temperature in deg R.

Samples from the cold ends of the bolts were tempered for various times in order to evaluate the relationship between the parameter P and hardness. Then P was determined from the observed hardness at any point along the bolt and the temperature at that point estimated for the service life of 17,500 hr. Fig. 4 shows the temperature estimated by this method as a function of the distance from the hot end of the bolt. During the later stages of this investigation, thermocouples were inserted in two of the replacement bolts in the throttle valves, and these actual temperature measurements are shown also in Fig. 4. The agreement between temperatures estimated from hardness measurement and that determined directly is satisfactory from about 960 to 850 F. Below 850 F the parameter method does not apply because the material had been tempered previously to the corresponding hardness.

The tensile properties of material at the free and fixed ends of one of the failed bolts were as follows:

	Fixed end	Free end	Specified
Tensile strength, psi	123250	137250	125000 min
Yield strength, 0.2 per cent offset, psi	99000	110000	105000 min
Elongation, per cent	17	18	16 min
Reduction of area, per cent	52.7	53.3	45 min

Both the tensile strength and yield strength at the fixed end of

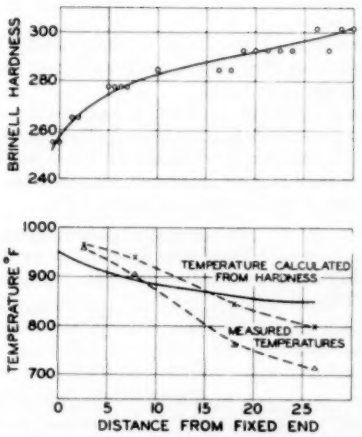


Fig. 4 Hardness and temperature distribution along bolts from Turbine A

the bolt were lower than those at the free end. The reduction in strength toward the fixed end of the bolt confirms the tempering during service previously indicated by hardness tests. The tensile properties at the free end of the bolt indicate that prior to service the material exceeded the specification requirements.

The failures of the bolts in turbine A occurred at a stress-concentration point at the root of the first engaged thread at the fixed ends of the bolts. Appreciable deformation of the bolt shanks occurred only toward the fixed ends of the bolts. Hardness measurements indicated that the fixed ends of the bolts were at a considerably higher temperature than were the free ends of the bolts. Direct temperature measurements confirmed the temperature distribution estimated from hardness measurements.

Bolting Failure, Turbine B

This 125,000-kw turbine, operating on 1000 F steam, had been in service for about one year. After this length of service, four bolts in a stop valve were found to have broken. Ultrasonic inspection indicated that additional bolts in this joint contained cracks.

These bolts also were provided with spacers to increase their effective elasticity. The spacers were shorter, approximately 8 in. long, than were those described previously for turbine A. During the first year of operation, this joint had been opened and the bolts were retightened after approximately 750, 4000, 4400, and 7300 hr. In each case the bolts were strained 0.0018 in. per in. or to 41,500 psi at temperature. The stop-valve bolts had been specified as a chromium-molybdenum-vanadium steel having considerably higher vanadium content than ASTM A 196, grade B 14 material. Data supplied by the material vendor and confirmed by tests in this laboratory indicated some improvement in relaxation resistance of this high-vanadium steel as compared to that of grade B 14.

As was the case of bolts from turbine A, these bolts failed with no visible ductility at the root of the first thread engaged in the valve body. Cracks in the bolts, which had not failed, also occurred at this location.

The chemical analyses of these bolts were as follows:

	Per cent		Per cent
Carbon	0.30	Molybdenum	0.54
Manganese	0.70	Silicon	0.69
Chromium	1.42	Vanadium	0.83

This analysis met the specification requirements for the special bolting material.

The deformation of the bolts, as determined by diameter measurements along their length, was greater near the fixed ends of the bolts as shown in Fig. 5.

There was no significant variation in hardness along the length of the chromium-molybdenum-high-vanadium bolts from the stop valve. Tensile properties at the fixed and free ends of one of these bolts were in good agreement with those determined in inspection tests of the material, as follows:

	After service		Before service	Specified
	Fixed end	Free end		
Tensile strength, psi	135500	136000	133000	125000
Yield strength, 0.2% offset, psi	118500	119000	116000	105000
Elongation, per cent	19	19.5	21	16
Reduction of area, per cent	57.5	58.3	64	45

Apparently, this higher vanadium modification of the B-14 material is resistant to tempering at the service temperature.

Failures of the bolts in turbine B were similar to those of bolts in turbine A. In both cases failures occurred at the root of the first engaged thread at the fixed ends of the bolts. This location is at an obvious stress concentration. Deformation of the shanks only toward the fixed ends of the bolts suggest a temperature gradient along the bolts. In the case of bolts from turbine B the temperature gradient could not be confirmed because the high-vanadium bolting material was resistant to tempering at the service temperature.

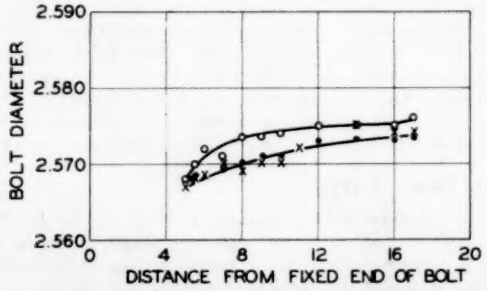


Fig. 5 Variation of diameter along bolt shank—bolts from Turbine B

Properties of Chromium-Molybdenum-Vanadium Bolting Material at Elevated Temperatures

The actual rupture of a material at a high temperature is dependent upon both time and stress. For many materials the relationship between the time to rupture and the initial stress in the material may be conveniently shown on logarithmic graph paper. Fig. 6 shows the available stress-rupture data for unnotched bars of B 14 bolting material and of the high-vanadium modification. This figure is based on data from the literature (2) and tests made in this laboratory. The rupture strength of the B 14 material at 900 F is very high even for long times. At higher temperatures the rupture strength decreases markedly especially for the longer test times. At 1000 F the rupture strength of the high-vanadium material is almost the same as that of the B 14 material, but at 1100 F the high-vanadium material is somewhat stronger. For both materials ductility decreases drastically, from 16 per cent to about three per cent elongation, as the time for rupture increases from about 10 to 2000 hr.

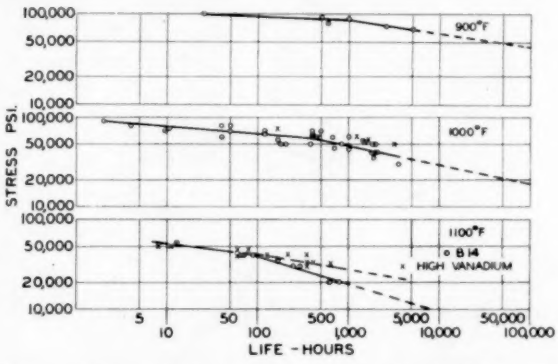


Fig. 6 Stress-rupture properties of chromium-molybdenum-vanadium bolting steel

Stress concentrations imposed by notches may modify greatly the stress-rupture behavior of materials. If the nominal stress at the root of the notch is compared with the nominal stress in a smooth test bar, a material may be either strengthened or weakened by the presence of a notch. Brown, Jones, and Newman have presented extensive data on notched rupture tests of the chromium-molybdenum-vanadium bolting material extending to about 1000 hr (3). Their data for notched and unnotched bars have been replotted in Fig. 7 and are compared to the average of the smooth-bar data presented previously. For their longest tests at

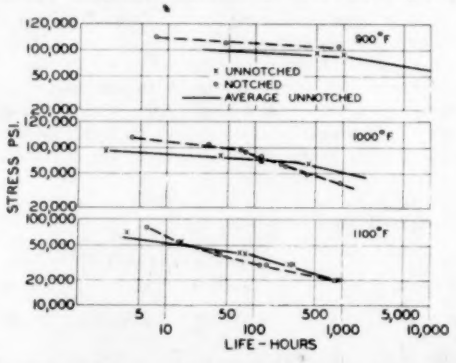


Fig. 7 Effect of sharp notches on rupture strength of chromium-molybdenum-vanadium bolting material

900 F the nominal stress for rupture of notched bars was considerably higher than that for smooth bars. At 1000 F, however, the stress for rupture of notched bars dropped to that of smooth bars at about 200 hr life and was less for longer times. Their data indicated a minimum notched-rupture strength about 65 per cent of the rupture strength of a smooth bar. At 1100 F the strength of the notched bars dropped below that of smooth bars at only 20 hr, then approached the smooth-bar strength at about 1000 hr. Probably, the notch sensitivity of the chromium-molybdenum-vanadium bolting steels increases to a maximum near 1000 F. Previous satisfactory experience with chromium-molybdenum-vanadium bolting below about 950 F is probably largely dependent upon the very high rupture strength combined with low notch sensitivity at the lower temperatures.

In bolting applications creep occurs under a condition of limited total strain, and the initial elastic extension of the bolt decreases during creep. Consequently, the stress in the bolt decreases during service. Robinson has described methods for the evaluation of the resistance of bolting materials to relaxation, or reduction of stress during service (4). Probably the most commonly used method is the step-down creep test with a fixed maximum strain in the specimen. For many materials, step-down creep-test data approximate the relationship

$$r = r_0 \left(\frac{S}{S_0}\right)^n$$

where r_0 is the creep rate at the stress S_0 and r is the creep rate at the stress S. If this relationship holds, the relaxation behavior of a bolt is approximately

$$S = \left[\frac{bS_0^n}{(n-1)r_0Et}\right]^{1/(n-1)}$$

where

r_0 = creep rate at stress S_0
b = elastic follow-up factor for bolted system
= 1 for a simple through bolt with a rigid flange
n = exponent of equation for creep rate
E = modulus of elasticity of bolt
S = stress at end of time t

This equation represents a straight line with a slope of $-1/(n-1)$ on a log-log graph of stress versus time. The values of n and S_0 are dependent upon the limiting strain in the step-down creep test. Hence this formula will not predict the relaxation behavior of bolts having initial strains significantly different from the limiting strain used in the step-down creep test. Available data from the literature (2, 4) and determined in this laboratory, on the relaxation behavior of chromium-molybdenum bolting materials are shown in Fig. 8, together with average rupture strengths taken from Fig. 6. At 900 F there is ample margin between the rupture strength and relaxed stress at all times. At 1000 F for long times the relaxed stress approaches the rupture strength of notched bars.

Because of the strain dependence of n and S_0, it is necessary to simulate elastic follow-up in the step-down test in order to obtain constants for exact analysis of a bolting system having elastic follow-up. Bolts having elastic follow-up have been considered in detail by Frey (5). He observed that calculated residual stresses of chromium-molybdenum-vanadium bolting steel, based upon constants determined with no elastic follow-up, were significantly lower, approximately 20 per cent, than the residual stresses experimentally determined with elastic follow-up. Relaxation tests, with and without elastic follow-up, on Type 422 stainless made in this laboratory are in reasonable agreement with Frey's observations for chromium-molybdenum-vanadium steel.

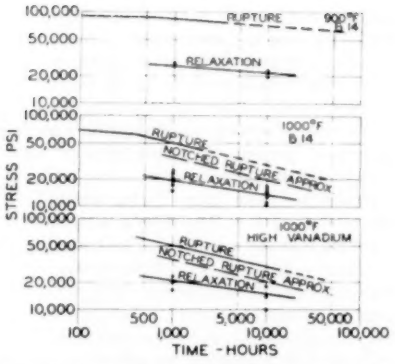


Fig. 8 Comparison of average stress-rupture data with available relaxation data for chromium-molybdenum-vanadium bolting materials

Prior plastic strain, particularly creep strain, also has a significant effect on the relaxation behavior of many metals. After completion of step-down relaxation tests on several samples of Type 422 steel, the samples were retested. In the second test, the material was much more resistant to relaxation, and the residual stresses were 8 to 44 per cent higher, average 25 per cent, than those observed in the first test. Data on chromium-molybdenum-vanadium bolting materials reported by Robinson (4) indicate relaxed stresses on restressing 6 to 31 per cent, average 16 per cent, higher than those obtained in the first test.

Analysis of Failures

Stress-rupture and relaxation data are not available at the actual bolt temperature, approximately 960 F. Hence one fundamental assumption to be made in analysis of these failures is that, over a small temperature range, rupture and relaxation behavior are proportionally affected by temperature. Material properties data at 1000 F will be used for analysis of the failures. A second assumption will be that the stress concentration caused by the threads causes a reduction in rupture strength of the bolts to 70 to 80 per cent of that of smooth bars.

The bolting of a turbine A was provided with 14-in. spacers, having cross-sectional areas equal to those of the bolts, in order to increase its elasticity. The flange through which the bolts passed was approximately 10 in. thick. If it is assumed that the portion of the bolt within the flange relaxes and the remainder of the bolt and spacer behave elastically, the follow-up factor $b = (28 + 10)/10 = 3.8$. Frey gave results of relaxation tests on chromium-molybdenum-vanadium bolting material for $b = 3.7$ and his data are replotted in Fig. 9 with average rupture data. A

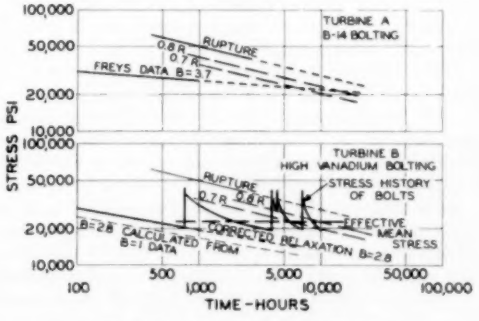


Fig. 9 Analysis of bolting failures

line extrapolated from Frey's data for $b = 3.7$ intersects lines representing 70 per cent and 80 per cent of the smooth-bar rupture strength at 7000 and 16,000 hr, respectively. These times are in reasonable agreement with the actual bolt life of 17,500 hr.

The analysis of failure of turbine B is not as direct as that for turbine A. No experimental data are available for relaxation with elastic follow-up of the chromium-molybdenum high-vanadium bolting material which had been used. Furthermore, the bolts had been retightened four times during service. These bolts had an elastic follow-up $b = 2.8$. Frey's data indicate that for $b = 2.8$ the relaxed stress calculated from constants determined at $b = 1$ is about 15 per cent low. Relaxed stresses of type 422 stainless steel calculated from constants determined at $b = 1$ were 3 to 31 per cent, average 13 per cent, lower than those determined for $b = 3$ in this laboratory. The relaxation behavior of the bolts in turbine B may be estimated by calculation of the relaxed stresses for $b = 2.8$ from constants determined at $b = 1$ and adding 15 per cent. This relaxation line for the bolts is shown as B in Fig. 9, with average rupture data. This line does not approach the notched-rupture strength of the material and the bolts should not fail. However, the bolts had been retightened and the actual stress history of the bolts was probably as shown by the heavy broken line. This line is drawn on the assumption that the relaxed stress after restressing is 20 per cent (an average of data for chromium-molybdenum-vanadium and type 422 steels) higher than the stress after the same time in the first stressing. The mean or effective stress for the approximately 3000-hr periods was calculated as

$$\bar{S} = 1.2 \frac{\int_1^{3000} S dt}{t} = 23{,}000 \text{ psi}$$

where S is defined by the equation of the corrected relaxation line $S = 65{,}000\, t^{-0.175}$, t is time, and 1.2 is the correction factor for restressing.

The mean stress line intersects lines representing 70 per cent and 80 per cent of the smooth-bar rupture strength at 6000 and 10,000 hr, respectively. These times are in reasonable agreement with the actual bolt life of 10,000 hr.

Summary and Conclusions

Analysis indicates that the failure of bolts from turbine A was caused by a combination of the following:

1 Stress concentration at the root of the first engaged thread.

2 Notch sensitivity in rupture of the chromium-molybdenum-vanadium material.

3 High relaxed stress imposed by elastic follow-up of bolt and spacer design combined with temperature gradient along the bolts.

These factors also were important in failure of bolts from turbine B. However, the elastic follow-up of these bolts alone was not sufficient to cause failure. Repeated restressing to a high level was the major factor in maintenance of a mean stress sufficient to cause failure. In both cases the stress concentration at the root of the threads reduced the strength of the bolt to about 80 per cent of the smooth-bar rupture strength. The maximum strain in the shanks of these bolts was only 0.6 to 1.2 per cent.

Bolted joints should be designed so that the mean relaxed stress will be well below the rupture strength of the bolts at the expected service life. The mean relaxed stress should be estimated based on elastic follow-up and expected restressing of the bolts during service. The smooth-bar rupture strength should be corrected by a suitable factor if the material is notch sensitive. For chromium-molybdenum-vanadium steels at about 1000 F the factor for sharp notches, 65 per cent, is recommended.

Because simple bolts without elastic follow-up relax rapidly, they should be safe if they are not restressed too frequently regardless of notch sensitivity of the material. Bolting with large elastic follow-up or which is frequently restressed may fail without warning if the material is notch sensitive. In these cases, a reduction in initial tightening stress, in order to reduce the mean stress, is recommended.

Materials which are not notch sensitive in rupture should be considered for bolting with follow-up or which must be retightened. Material of this type will deform excessively before rupture and thus provide warning of overstress.

Bibliography

1 "Time Temperature Relations in Tempering Steels," by J. H. Holloman and L. D. Jaffe, *AIME Trans.*, Iron and Steel Division, vol. 162, 1945, pp. 223–249.

2 "Medium Carbon Pearlitic Alloy Steels for High Temperature Applications," Timken Roller Bearing Company Technical Bulletin 36A.

3 "Influence of Sharp Notches on the Stress Rupture Characteristics of Several Heat Resisting Alloys," by W. F. Brown, M. H. Jones, and D. P. Newman, American Society for Testing Materials, Special Technical Publication No. 128, 1953, pp. 25–45.

4 "High Temperature Bolting Materials," by E. L. Robinson, *ASTM Proceedings*, vol. 48, 1948, pp. 214–235.

5 "The General Tensional Relaxation Properties of a Bolting Steel," by D. N. Frey, TRANS. ASME, vol. 73, 1951, pp. 755–760.

The Heat-Balance Integral and Its Application to Problems Involving a Change of Phase[1]

By T. R. GOODMAN,[2] BOSTON, MASS.

An approximate mathematical technique utilizing the "heat-balance integral" is presented for solving for the location of the melt line in heat-conduction problems involving a change of phase. Analytical expressions are derived when (*a*) boundary temperature is fixed; (*b*) heat flux at boundary is given; (*c*) heat flux is generated aerodynamically or by radiation; (*d*) heat flux at boundary is given and melt is completely removed; (*e*) heat flux at boundary is given, and at time t_0 melt begins to vaporize. Comparisons with known solutions have been made when available, and ultimately all the solutions are presented in graphical form.

I Introduction

THE heat-conduction problem involving a change of phase (sometimes called the problem of Stefan) is nonlinear because it involves a moving boundary (the melt line) whose location is unknown *a priori*. Except for one special case (1),[3] no analytical technique exists for finding solutions. Evans, Isaacson, and MacDonald (2) have presented some solutions for the location of the melt line in the form of a Taylor series in time, the convergence of which is undetermined. Landau (3) has presented another set of solutions obtained by solving the heat-conduction equation numerically with a finite-difference approximation. The intricate calculations required for a finite-difference solution with a change of phase are presented in detail by Forster (4).

The class of problems characterized as nonlinear heat-transfer problems, of which those involving a change of phase are a subclass, presents formidable mathematical difficulties. The finite-difference procedure is tedious, and what is more, it must be repeated each time a parameter is changed. Recently two methods have been presented which may be used in nonlinear cases to obtain approximate solutions. J. W. Green (5) has applied Galerkin's method to the heat-transfer problem, and although linear problems only are discussed, it is clear that the method is equally applicable to those which are nonlinear. M. A. Biot (6) recently published a method for solving nonlinear heat-transfer problems, and applied it to those problems whose nonlinearities arise because of temperature-dependent transport properties.

This paper presents a method which utilizes the "heat-balance integral," and applies it to problems involving a change of phase. For the case of one space variable (the only case considered) the equation for determining the location of the moving boundary reduces to an ordinary differential equation which, for many cases of interest, can be solved analytically. In particular, the solutions in references (1, 2, 3) are reworked for purposes of comparison. In addition, two new problems are solved; viz., melting due to aerodynamic heating or radiation, and the vaporization of a melting solid.

II Heat-Balance Integral and Thermal Layer

As an introduction to the use of the heat-balance integral a simple heat-conduction problem, which does not involve a change of phase, will be solved. The exact solution appears in Carslaw and Jaeger (1) and will be used for comparison.

Assume a semi-infinite slab extending over positive x. Initially, the temperature u is $-V$, and at the surface $x = 0$ the heat flux $H(t)$ is given for time $t > 0$. If κ is the thermal diffusivity, the heat-conduction equation is

$$\kappa \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} \qquad x > 0, \quad t > 0 \ldots\ldots\ldots [1]$$

If k is the thermal conductivity, the boundary condition is

$$k \frac{\partial u}{\partial x} = -H(t) \qquad x = 0, \quad t > 0 \ldots\ldots\ldots [2]$$

A quantity $\delta(t)$ is now defined to be the thermal layer. For $x > \delta(t)$, the slab, for all practical purposes, is at an equilibrium temperature and there is no heat transfer beyond this point. $\delta(t)$ is completely equivalent to the boundary-layer thickness in hydrodynamics. If Equation [1] is now multiplied by dx and integrated from 0 to δ, the resulting equation will be called the heat-balance integral. The heat-conduction equation will, thereby, be satisfied only on the average. This averaged equation is analogous to the momentum integral in boundary-layer theory. The method of the momentum integral is due originally to von Karman and Pohlhausen (7). A modern account of the Karman-Pohlhausen method and a bibliography may be found in reference (8). Although the general technique was introduced by Karman and Pohlhausen to solve boundary-layer problems in fluid mechanics, it is nevertheless applicable to the solution of all problems governed by a diffusion-type equation. Such problems as the nonsteady heat transfer in a solid, and the nonsteady flow of fluids through porous media are of this type. If the problems are linear they usually can be solved by exact methods. Heat-transfer problems involving a change of phase, on the other hand, are nonlinear and, except in very special cases, must be solved either by integrating the heat-conduction equation numerically, or by using some approximate technique. The method of the heat-balance integral will be developed, and it will be shown how the method yields simple analytical solutions for heat-conduction problems involving a change of phase. These solutions, although not exact, have accuracy which is useful to the engineer.

[1] This work was performed under Air Force Contract No. 04(645)-24 for the Special Defense Projects Department of the General Electric Company.

[2] Senior Engineer, Allied Research Associates, Inc.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the Heat Transfer and Fluid Mechanics Institute, Pasadena, Calif., June 21, 1957.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 21, 1957.

The heat-balance integral obtained by averaging Equation [1] in the manner described becomes

$$\frac{d}{dt}(\theta + V\delta) = \kappa\left[\frac{\partial u}{\partial x}(\delta, t) - \frac{\partial u}{\partial x}(0, t)\right] \quad [3]$$

where

$$\theta = \int_0^{\delta(t)} u dx \quad [4]$$

But, since there is no heat transferred beyond $x = \delta$

$$\frac{\partial u}{\partial x}(\delta, t) = 0 \quad [5]$$

Assume that u can be represented by a second-degree polynominal in x in the form, $u = a + bx + cx^2$ where the coefficients may depend on t. Applying Equations [2], [5] and the condition that at $x = \delta$, $u = -V$, the following equation results

$$u = -V + \frac{H}{2k\delta}(\delta - x)^2 \quad [6]$$

By virtue of Equation [4]

$$\theta = -V\delta + \frac{\delta^2 H}{6k} \quad [7]$$

Introducing Equations [2], [5], and [7] into the heat-balance integral, Equation [3] gives

$$\frac{1}{6}\frac{d}{dt}(\delta^2 H) = \kappa H \quad [8]$$

By virtue of the initial condition $\delta(0) = 0$

$$\delta = \sqrt{(6\kappa)}\left[\frac{1}{H(t)}\int_0^t H(t_1)dt_1\right]^{1/2} \quad [9]$$

If $H(t)$ is constant

$$\delta = \sqrt{(6\kappa t)} \quad [10]$$

The surface temperature is obtained by setting $x = 0$ in Equation [6] and applying Equation [9]. The result is

$$u(0, t) = -V + \sqrt{(3/2)}\sqrt{\kappa}\left[H(t)\int_0^t H(t_1)dt_1\right]^{1/2}\Big/k \quad [11]$$

If $H(t)$ is constant

$$u(0, t) = -V + \sqrt{(3/2)}\sqrt{\kappa}\,H\sqrt{(t)}/k \quad [12]$$

The surface temperature for this problem is known exactly (1) and the result is for a constant value of H

$$u(0, t) = -V + \sqrt{(4/\pi)}\sqrt{\kappa}\,H\sqrt{(t)}/k \quad [13]$$

A comparison of Equations [12] and [13] discloses the fact that the results are of the same form, and differ only by a numerical factor. Since $\sqrt{(4/\pi)} = 1.13$ and $\sqrt{(3/2)} = 1.23$, the error is about 9 per cent. This error can be reduced to 2 per cent by using a cubic to represent the temperature.[4]

It will be required for subsequent development to know the value of δ when $u(0, t) = 0$. From Equations [9] and [11] the result is

$$\delta_m = 2Vk/H(t_m) \quad [14]$$

where t_m is the time at which $u(0, t) = 0$ occurs, and can be obtained from Equation [11] or [12]. From Equation [12] the result is

$$t_m = \frac{2}{3}\frac{V^2k^2}{\kappa H^2} \quad [15]$$

III Melting of a Solid With Fixed Boundary Temperature[5]

Solution Using the Heat-Balance Integral. The method of the heat-balance integral, although applicable to all heat-transfer problems, finds its greatest utility in those involving a change of phase or having nonlinear boundary conditions. Analytic methods do not exist in these cases, and one must resort either to numerical methods or to approximate solutions. The problem of a melting solid with a fixed boundary temperature can be solved exactly by use of a similarity variable, $\eta = x/\sqrt{(\kappa t)}$, and the exact solution will serve as a check on the approximate one to be derived.

Define $x = s(t)$ to be the location of the melt line. The melting temperature will be chosen to be zero for this and all subsequent problems. Hence

$$u[s(t), t] = 0 \quad [16]$$

Furthermore, from a heat balance across the melt line

$$k_-\frac{\partial u}{\partial x}(s_-, t) - k_+\frac{\partial u}{\partial x}(s_+, t) = \pm\rho L\frac{ds}{dt} \quad [17]$$

where ρ is the density and L the latent heat of melting. This equation states that the heat flux entering the melt line minus the flux leaving the melt line equals the latent heat absorbed or emitted depending on whether the problem is one of freezing or melting.

The problem will now be simplified by assuming all the solid to be at the melting temperature. The thermal layer is then compelled to be identical to the melt layer; i.e., $\delta = s$. Actually, of course, there is a temperature distribution in the solid as well. Although the present method is capable of taking this into account, the equations are considerably more complicated. The location of the melt line depends on the boundary temperature, and also on the temperature as $x \to \infty$. It is shown in Carslaw and Jaeger, however, that the dependence on the temperature as $x \to \infty$ is small in comparison to the dependence on the boundary temperature. Even if the temperature as $x \to \infty$ is considerably smaller than the melting temperature, the solid will have been heated to some extent before the melting temperature is reached on the boundary. If this heating has been slow the temperature gradient in the solid will be small, and the solid will be virtually at the melting temperature in the neighborhood of the melt line. Therefore the temperature distribution in the solid will not affect the location of the melt line to any great extent. For a mathematical discussion of this point see Evans, Isaacson, and MacDonald (2).

With the assumption of constant temperature in the solid, Equation [17] reduces to

$$\frac{\partial u}{\partial x}(s, t) = -A\frac{ds}{dt} \quad [18]$$

where $A = \rho L/k$, and k is the thermal conductivity of the liquid.

The boundary condition is

$$u = V_1, \quad x = 0, \quad t > 0 \quad [19]$$

[4] If the slab were of finite thickness l, then it would behave as a semi-infinite slab until $\delta = l$ after which time, according to this theory, δ would always be l. Equation [5] would be replaced by some given condition at $x = l$. The condition that $u = -V$ at $x = \delta$ would be abandoned, thus leaving one of the constants in the quadratic representation of the temperature distribution unknown. Finally, this constant would be calculated from the heat-balance integral, the initial value being determined from the semi-infinite slab solution when $\delta = l$.

[5] The solution is equally applicable to the freezing of a liquid.

Equations [1], [16], [18], [19] constitute a complete statement of the problem.

The heat-balance integral becomes

$$\frac{d\theta}{dt} = -\kappa\left[A\,\frac{ds}{dt} + \frac{\partial u}{\partial x}\,(0, t)\right] \quad [20]$$

where, θ is essentially the total energy in the melt, and is given by

$$\theta = \int_0^{s(t)} u(x, t)dx \quad [21]$$

Once again let u be represented by a second-degree polynomial in x. Three conditions are necessary in order to obtain the constants. Equations [16] and [19] are two conditions, and the third condition is essentially Equation [18]. But in its present form Equation [18] is not suitable because the coefficients in the polynomial would involve ds/dt. In turn, θ would involve ds/dt, and the heat-balance integral would then become a second-order differential equation for $s(t)$ whereas there is only one initial condition for s; namely, $s(0) = 0$. To circumvent this difficulty, differentiate Equation [16] with respect to t

$$\frac{\partial u}{\partial x}\,\frac{ds}{dt} + \frac{\partial u}{\partial t} = 0 \quad [22]$$

If this is solved for ds/dt and the result substituted into Equation [18], there is obtained

$$\left(\frac{\partial u}{\partial x}\right)^2 = A\,\frac{\partial u}{\partial t} \quad [23]$$

But a partial derivative with respect to time is inadmissible for determining the constants in the polynomial, because the constants would then be determined from a differential rather than an algebraic equation. Substitute, therefore, from Equation [1] for $\partial u/\partial t$. The third condition is then seen to be

$$\left(\frac{\partial u}{\partial x}\right)^2 = \kappa A\,\frac{\partial^2 u}{\partial x^2}, \quad x = s \quad [24]$$

With the boundary condition in this form the nonlinearity of the problem becomes self-evident.

If the temperature distribution is given by

$$u = a(x - s) + b(x - s)^2 \quad [25]$$

then the quantities a and b are determined by

$$a = \frac{A\kappa}{s}\,[1 - (1 + \mu)^{1/2}] \quad [26]$$

$$b = \frac{as + V_1}{s^2}$$

where

$$\mu = 2V_1/A\kappa \quad [27]$$

The quantities θ and $[\partial u/\partial x(0, t)]$ are readily determined, and the heat-balance integral, Equation [20], ultimately yields the following differential equation for s

$$s\,\frac{ds}{dt} = \frac{6\kappa[1 - (1 + \mu)^{1/2} + \mu]}{5 + \mu + (1 + \mu)^{1/2}} \quad [28]$$

If the initial condition $s(0) = 0$ is applied, the solution is

$$s = K\sqrt{t} \quad [29]$$

where

$$\frac{K}{2\sqrt{\kappa}} = \sqrt{3}\left[\frac{1 - (1 + \mu)^{1/2} + \mu}{5 + (1 + \mu)^{1/2} + \mu}\right]^{1/2} \quad [30]$$

Solution Using Boundary Condition at Melt Line. Once the temperature distribution was determined from Equations [25], [26], [27], the differential equation for s was derived from the heat-balance integral. But there exists an alternate possibility; if $[\partial u/\partial x(s, t)]$ obtained from the known distribution is substituted into the boundary condition at the melt line, Equation [18], a differential equation for s which is different from Equation [28] is the result. This equation is

$$s\,\frac{ds}{dt} = \kappa[-1 + (1 + \mu)^{1/2}] \quad [31]$$

The solution is of the same form as Equation [29] except that the constant K is given by

$$\frac{K}{2\sqrt{\kappa}} = \left[\frac{(1 + \mu)^{1/2} - 1}{2}\right]^{1/2} \quad [32]$$

The question naturally arises, which of the two values for K, that given by Equation [30] or that given by [32], should be used? Only a comparison with the exact result will determine which is more accurate. The exact solution is given in Carslaw and Jaeger (1)

$$\mu = 2\sqrt{\pi}\left(\frac{K}{2\sqrt{\kappa}}\right)\exp\left(\frac{K}{2\sqrt{\kappa}}\right)^2 \operatorname{erf}\left(\frac{K}{2\sqrt{\kappa}}\right) \quad [33]$$

The three Equations, [30], [32], and [33], are shown plotted in Fig. 1 for purposes of comparison, and it is seen that Equation [32] is the more accurate. It would be desirable to avoid the necessity for choosing between the differential equation which arises from the heat-balance integral, and that which arises from the boundary condition at the melt line. In order to do this, the temperature distribution may be taken to be a cubic, thereby introducing another unknown, c. The two differential equations may then be solved simultaneously for s and the additional unknown. It would be expected that the result obtained using a cubic temperature distribution would be more accurate than that which used a quadratic, and this is borne out as seen in Fig. 1.

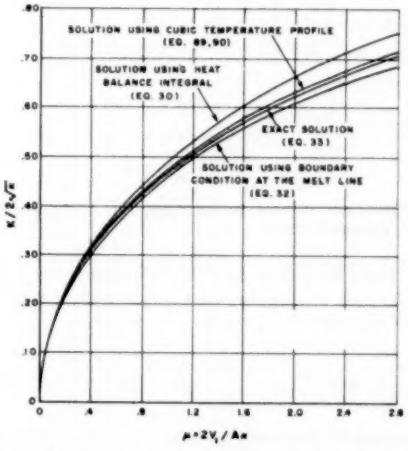


Fig. 1 Melting constant for fixed boundary temperature

The analysis is presented in the Appendix, and it is seen that the solution is considerably more complicated than either Equation [30] or [32].

IV Melting of Solid With Given Heat Flux at Boundary[6]

All the solid is once again assumed to be at the melting tem-

[6] The solution is equally applicable to the freezing of a liquid.

perature. The problem is identical to the preceding one except that the condition at the boundary, Equation [19], is replaced by Equation [2]. The heat-balance integral becomes

$$\frac{d}{dt}(\theta + \kappa A s) = \kappa H(t)/k \quad [34]$$

Integrating and applying the initial condition, $s(0) = 0$, there is obtained

$$\theta + \kappa A s = \frac{\kappa}{k}\int_0^t H(t_1)dt_1 \quad [35]$$

To determine θ, first assume u in the form of a second-order polynomial. The three conditions for determining the constants are Equations [2], [16], and [24]. θ is defined by Equation [21]. Substituting into Equation [35], the final result is

$$\tau = \frac{\sigma}{6}[\sigma + 5 + (1 + 4\sigma)^{1/2}] \quad [36]$$

where

$$\tau = \frac{H(t)}{\kappa k^2 A^2}\int_0^t H(t_1)dt_1 \quad [37]$$

$$\sigma = \frac{H(t)s}{\kappa A k} \quad [38]$$

Equation [36] is plotted in Fig. 2. The exact solution given by Evans, Isaacson, and MacDonald (2) is expressed in the form of

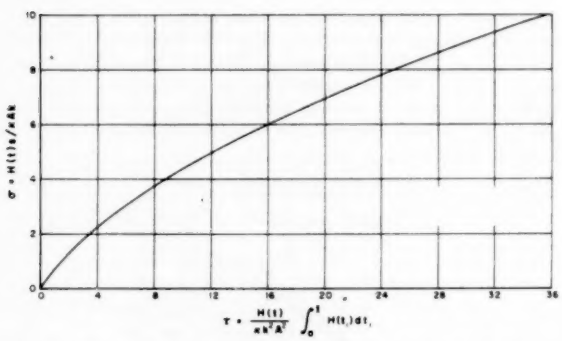


Fig. 2 Thickness of melt versus time, for a given heat flux at boundary, Equation [36]

a Taylor series for σ in terms of τ. Equation [36] must be expressed in terms of a Taylor series also in order to effect a comparison. First find the series for τ in terms of σ

$$\tau = \sigma + \frac{1}{2}\sigma^2 - \frac{1}{3}\sigma^3 + \frac{2}{3}\sigma^4 - \frac{5}{3}\sigma^5 + \ldots \quad [39]$$

This series can be inverted

$$\sigma = \tau - \frac{\tau^2}{2!} + \frac{5\tau^3}{3!} + \frac{51\tau^4}{4!} + \frac{795\tau^5}{5!} + \ldots \quad [40]$$

which may be compared with the exact series for the case H = const

$$\sigma = \tau - \frac{\tau^2}{2!} + \frac{5\tau^3}{3!} - \frac{51\tau^4}{4!} + \frac{827\tau^5}{5!} + \ldots \quad [41]$$

There is an error of about 4 per cent in the coefficient of τ^5.

The temperature-time history on the boundary is obviously of interest and the result is

$$\frac{u(0,t)}{A\kappa} = -\frac{1}{4} + \frac{1}{4}(1 + 4\sigma)^{1/2} + \sigma/2 \quad [42]$$

This can be cross-plotted with Equation [36] to find $u(0, t)/A\kappa$ in terms of τ, and the result is shown in Fig. 3.

Equations [36], [37], and [38] represent an approximate solution for an arbitrary heat input. But it has been discovered that it is useful only for functions $H(t)$ which are monotonically increasing or constant. If H is a pulse-type function which vanishes after some finite time, then according to Equation [38] σ vanishes when H vanishes, and hence, by virtue of Equation [42], the surface temperature also vanishes at the same time. In actuality this cannot, of course, occur, but instead the surface temperature will decay gradually after heat shutoff. It is thus seen that the solution fails sometime after H_{max} and breaks down completely at heat shutoff. This failure is reminiscent of the failure of the Karman-Pohlhausen one-parameter method in an adverse pressure gradient as one nears the separation point. The method of the heat-balance integral can still be used for pulse heat inputs by using a two-parameter method. If refreezing is to occur, it is necessary to take into account the temperature distribution in the solid and use a three-parameter method. These problems are, however, beyond the scope of the present paper.

V Melting of a Solid Due to Aerodynamic Heating or Radiation[7]

In this case the boundary condition is replaced by

$$k\frac{\partial u}{\partial x} = h[u - u_0], \quad x = 0, \quad t > 0 \quad [43]$$

If this boundary condition is interpreted as aerodynamic heating, then h is the heat-transfer coefficient and u_0 is a reference temperature of the external flow. If the boundary condition is interpreted as radiation, then h is the so-called exterior conductivity, u_0 is the temperature of the surrounding medium, and $(u - u_0/u_0) \ll 1$.

The method of solution by the heat-balance integral is similar to that of the preceding problems, although considerably more complicated. The final result is

$$T = \frac{1}{12\beta}\Bigg[[(1 + 2\beta) + (2 + \beta)S][1 + \beta S(2 + S)]^{1/2} - \frac{2(\beta - 1)}{\sqrt{\beta}}\ln\left\{\frac{[1 + \beta S(2 + S)]^{1/2} + [(1 + S)\beta]^{1/2}}{1 + \sqrt{\beta}}\right\} - 4\beta(\beta - 1)\ln\left\{\frac{-1 + \beta(2 + S) + [1 + \beta S(2 + S)]^{1/2}}{2\beta}\right\} + (\beta^2 + 5\beta)\frac{S^2}{2} + 2(\beta^2 + 4\beta - 2)S - (1 + 2\beta)\Bigg] \quad [44]$$

where

$$T = \frac{h^2 u_0 t}{k^2 A} \quad [45]$$

$$S = \frac{hs}{k} \quad [46]$$

$$\beta = 1 + \frac{2u_0}{A\kappa} \quad [47]$$

For the special case $\beta \to 1$ or infinite latent heat, Equation [44] reduces to

[7] The solution is equally applicable to the freezing of a liquid due to aerodynamic cooling.

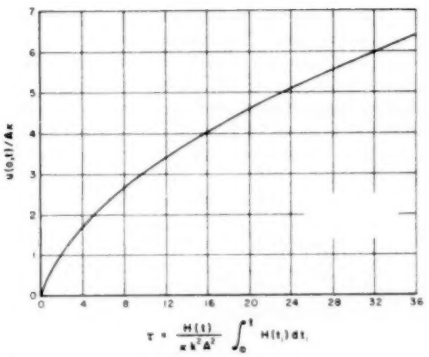


Fig. 3 Temperature-time history on boundary for a given heat flux at boundary, Equations [42] and [36]

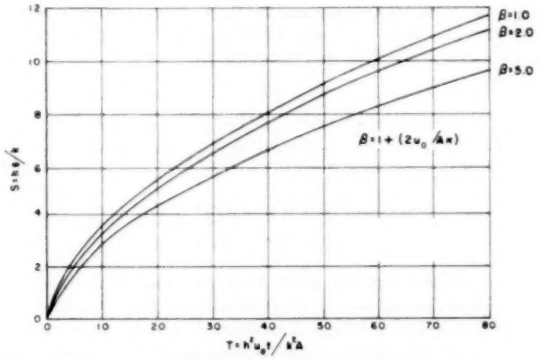


Fig. 4 Thickness of melt versus time, for aerodynamic heating or radiation boundary condition, Equation [44]

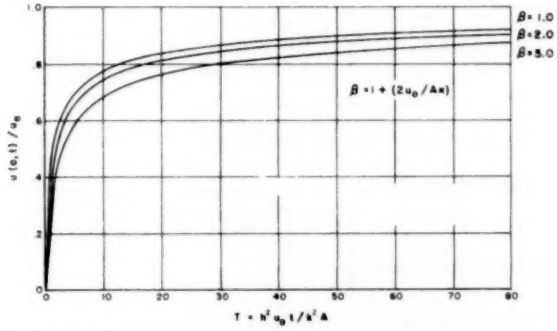


Fig. 5 Surface temperature versus time, for aerodynamic heating or radiation boundary condition, Equations [49] and [44]

$$T = \frac{S^2}{2} + S \quad \ldots\ldots [48]$$

A plot of S versus T for three values of β is shown in Fig. 4. The temperature-time history on the boundary is

$$\frac{u(0,t)}{u_0} = \frac{\{(\beta-1)S^2 + 2(\beta-2)S - 2 + 2[1+\beta S(2+S)]^{1/2}\}}{(\beta-1)(2+S)^2} \quad \ldots [49]$$

For $\beta \to 1$ Equation [49] reduces to

$$\frac{u(0,t)}{u_0} = \frac{S}{1+S} \quad \ldots\ldots [50]$$

Equations [49] and [50] have been cross-plotted with Equations [44] and [48] and the results are shown in Fig. 5.

VI Melting of a Solid With Complete Removal of Melt

In the general problem of a melting solid there is a temperature distribution in both the liquid and the solid. In the preceding problems, a vast simplification was achieved by assuming that the solid was at the (constant) melting temperature, and that a temperature distribution existed only in the liquid. In the present problem all the melt is immediately removed, and hence a temperature distribution exists only in the solid. Therefore this problem may be considered to be the inverse of the preceding ones, and simplifications of a different sort will be achieved.

It is assumed that the semi-infinite solid slab has been heated by application of a constant heat flux H at the boundary $x = 0$. At time $t = 0$ the melting temperature $u = 0$ is reached on the boundary, and at that time the thermal layer is given by Equation [14]. For positive time the solid melts, and all the melt is immediately swept away by some undisclosed mechanism (perhaps aerodynamic shearing forces). The boundary and melt line are now both located at $x = s(t)$, and the boundary condition from Equation [17] becomes[8]

$$H + k\frac{\partial u}{\partial x} = \rho L \frac{ds}{dt}, \quad x = s(t), \quad t > 0 \quad \ldots\ldots [51]$$

The temperature distribution in the solid is again represented by a second-degree polynomial satisfying the conditions: At $x = s$, $u = 0$; at $x = \delta$, $u = -V$, $(\partial u/\partial x) = 0$. This leads to a distribution of the form

$$u = V\left[-2\left(\frac{x-s}{\delta-s}\right) + \left(\frac{x-s}{\delta-s}\right)^2\right] \quad \ldots\ldots [52]$$

To obtain the heat-balance integral, the heat-conduction equation is averaged from s to δ. Applying Equation [51] to determine $\partial u/\partial x = 0$ at δ, the heat-balance integral becomes

$$\frac{d}{dt}\left(\theta + V\delta + \frac{\rho L \kappa s}{k}\right) = \frac{\kappa H}{k} \quad \ldots\ldots [53]$$

where, for this case

$$\theta = \int_s^\delta u\,dx \quad \ldots\ldots [54]$$

θ is easily determined from Equation [52], and then Equation [53] becomes

$$\frac{\kappa H}{Vk} = \frac{d}{dt}\left[\frac{\delta - s}{3}\right] + \frac{ds}{dt}(1+\nu) \quad \ldots\ldots [55]$$

where

$$\nu = \kappa\rho L/kV \quad \ldots\ldots [56]$$

Another differential equation is found from the condition at the melt line, Equation [51]. Substituting for $\partial u/\partial x$ from Equation [52] there is found

$$\rho L \frac{ds}{dt} = H - \frac{2kV}{\delta - s} \quad \ldots\ldots [57]$$

Equations [55] and [57] are two simultaneous differential equa-

[8] It is worth noting that, since the boundary temperature is constant, Equation [51] is the correct boundary condition when the heat is generated aerodynamically or by radiation.

tions for s and $\delta - s$. The initial conditions are $s(0) = 0$, $\delta(0) = 2Vk/H$ (see Equation [14]).

Assume that this pair of equations possesses a steady-state solution; i.e., assume ds/dt has a constant value of q. It then follows from Equation [57] that $\delta - s$ is constant, and from Equation [55] it is seen that

$$q = \frac{\kappa H/Vk}{1 + \nu} \quad [58]$$

This value of q is precisely that which Landau (3) obtained using the exact system of equations.

To solve the equations completely, eliminate ds/dt between them, and let

$$\zeta = \frac{H(\delta - s)}{Vk} \quad [59]$$

There remains a differential equation of ζ in terms of time. Let

$$\Omega = \frac{H^2 t}{\rho L k V} \quad [60]$$

The solution, after applying the initial conditions, is

$$\Omega = -\frac{1}{3}\left[\zeta - 2 + 2(1 + \nu) \ln \frac{2(1 + \nu) - \zeta}{2\nu}\right] \quad [61]$$

Substituting this into Equation [57] and defining

$$\mathbb{S} = \frac{sH}{Vk} \quad [62]$$

there is obtained

$$\mathbb{S} = -\frac{1}{3}\left[\zeta - 2 + 2\nu \ln \frac{2(1 + \nu) - \zeta}{2\nu}\right] \quad [63]$$

Equations [61] and [63] are the equations for the melt line in parametric form. They are plotted in Fig. 6 for a few values of ν. To the scale which the graph is plotted, the present results and those of Landau (3), which were obtained numerically, are indistinguishable.

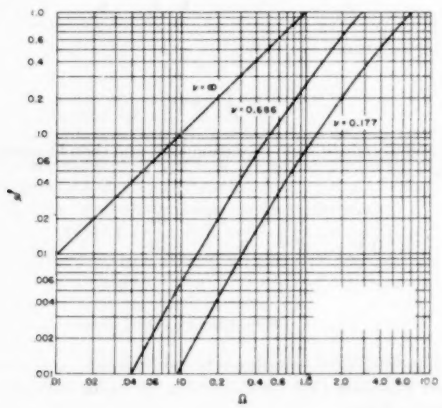


Fig. 6 Melt-line location versus time, with complete removal of melt, Equations [60] and [62]

Vaporization of a Melting Solid

It will be assumed that melting occurs with no removal of melt, according to the solution given by Equations [36], [37], [38], [42] At time t_0 the vaporization temperature V is reached on the boundary, and from that time onward the vapor is removed by forced or free convection. The location of the melt line will be denoted by s_1, and at time t_0, $s_1 = s_0$. The boundary condition at $x = s_1$ is given by Equation [18]. The location of the vaporization line will be denoted by s_2, and the boundary condition at the vaporization line is given by Equation [51] where L in this case is the latent heat of vaporization, and will be denoted subsequently by L_V.

Exact Steady-State Solution. Assume there is a steady-state solution, in which case $s_2 = qt$, $s_1 = qt + \bar{s}$, where q and $\bar{s}$ are constants to be determined. The temperature u in the interval $s_2 < x < s_1$ obeys the heat-conduction Equation [1] but, instead of depending on x and t separately, depends only on $(x - qt)$]. The general solution then becomes

$$u = Ae^{-\left[\frac{q}{\kappa}(x - qt)\right]} + B \quad [64]$$

where A and B are two more constants to be determined. The four conditions for determining the four constants are Equations [16] and [18] at $x = s_1$, and Equations [51] and the condition $u = V$ at $x = s_2$. The four constants become

$$q = \frac{2H/Ak}{2 + \mu(1 + \nu)} \quad [65]$$

$$\bar{s} = \frac{\kappa}{q} \ln (1 + \mu/2) \quad [66]$$

$$A = V/(1 - e^{-q\bar{s}/\kappa}) \quad [67]$$

$$B = -Ve^{-q\bar{s}/\kappa}/(1 - e^{-q\bar{s}/\kappa}) \quad [68]$$

where μ and ν are defined by Equations [27] and [56], respectively, except that L is replaced by L_V in Equation [56].

Complete Solution Using Heat-Balance Integral. Let the temperature be represented by a second-degree polynomial. The three conditions for determining the constants are: At $x = s_2$, $u = V$; at $x = s_1$, Equations [16] and [24] are valid. The temperature distribution is given by

$$u = V\left[1 - z_0\left(\frac{x - s_2}{s_1 - s_2}\right) - (1 - z_0)\left(\frac{x - s_2}{s_1 - s_2}\right)^2\right] \quad [69]$$

where

$$z_0 = \frac{2(1 + \mu)^{1/2}}{\mu}[(1 + \mu)^{1/2} - 1] \quad [70]$$

After defining θ to be

$$\theta = \int_{s_2}^{s_1} u\, dx \quad [71]$$

the heat-balance integral becomes

$$\frac{V}{6}(4 - z_0)\frac{d}{dt}(s_1 - s_2) + V\frac{ds_2}{dt} = \kappa\left[-A\frac{d}{dt}(s_1 - s_2) - \left(A + \frac{\rho L}{\kappa}\right)\frac{ds_2}{dt} + \frac{H}{\kappa}\right] \quad [72]$$

The condition at the vaporization line yields

$$H - \frac{kVz_0}{s_1 - s_2} = \rho L\frac{ds_2}{dt} \quad [73]$$

Equations [72] and [73] represent two simultaneous equations for determining s_2 and $(s_1 - s_2)$. They possess a steady-state solution which yields a value for q identical to Equation [65]. The value of $\bar{s}$, however, is given by

$$\bar{s} = \frac{\kappa}{q}\frac{(1 + \mu)^{1/2}[(1 + \mu)^{1/2} - 1]}{1 + \mu/2} \quad [74]$$

for small μ this may be expanded in a power series

$$\bar{s} = \frac{\kappa}{q}\frac{\mu}{2}\left[1 - \frac{\mu}{4} + 0(\mu^2)\right] \quad [75]$$

The exact value of $\bar{s}$ represented by Equation [66] may be expanded to yield

$$\bar{s} = \frac{\kappa}{q}\frac{\mu}{2}\left[1 - \frac{\mu}{4} + \frac{\mu^2}{12} + \ldots\right] \quad [76]$$

The accuracy of Equation [75] is indicative of the accuracy of the complete solution.

Before proceeding to the complete solution of Equations [72] and [73], define a dimensionless time by Equation [60]. The dimensionless value of the thickness of melt is defined by

$$z = \frac{H(s_1 - s_2)}{kV} \quad [77]$$

The dimensionless value for the location of the vaporization line is

$$\Gamma = \frac{s_2 H}{Vk} \quad [78]$$

The initial value of s_2 is $s_2(t_0) = 0$. The initial value of s_1 may be obtained from Equation [42] by setting $u(0, t) = V$. The result is

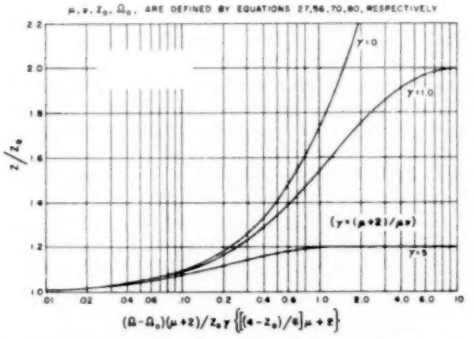


Fig. 7 Thickness of melt versus time, for a melting solid which is vaporizing, Equation [81]

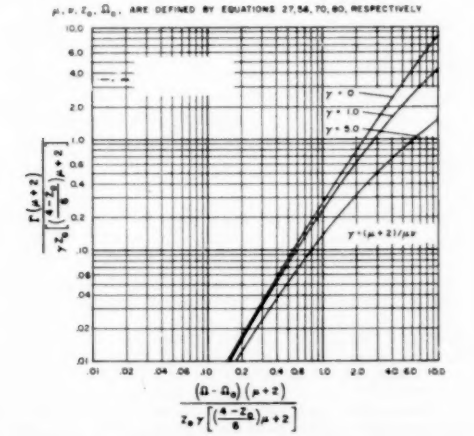


Fig. 8 Location of vaporization line versus time, for melting solid which is vaporizing, Equations [81] and [82]

$$z(t_0) = z_0 \quad [79]$$

where z_0 is defined by Equation [70]. Furthermore, by virtue of Equation [36]

$$\Omega_0 \equiv \frac{H^2 l_0}{\rho L V k} = \frac{z_0}{3\mu\nu}\left[\frac{\mu z_0}{2} + 5 + (1 + 2\mu z_0)^{1/2}\right] \quad [80]$$

The complete solution is

$$\Omega - \Omega_0 = \frac{z_0}{\gamma(\mu+2)}\left[\left(\frac{4 - z_0}{6}\right)\mu + 2\right]\{(1 - z/z_0)\gamma - (1+\gamma)\ln[1 + \gamma(1 - z/z_0)]\} \quad [81]$$

$$\Gamma = \frac{z_0}{\gamma(\mu+2)}\left[\left(\frac{4 - z_0}{6}\right)\mu + 2\right]\{(1 - z/z_0)\gamma - \ln[1 + \gamma(1 - z/z_0)]\} \quad [82]$$

where

$$\gamma = \frac{\mu+2}{\mu\nu} = \frac{H}{\rho q L} - 1 \quad [83]$$

For $\nu \to \infty$ with μ fixed

$$\Omega - \Omega_0 = \frac{z_0\gamma}{2(\mu+2)}\left[\left(\frac{4 - z_0}{6}\right)\mu + 2\right]\left[\left(\frac{z}{z_0}\right)^2 - 1\right] \quad [84]$$

$$\Gamma = \frac{z_0\gamma}{2(\mu+2)}\left[\left(\frac{4 - z_0}{6}\right)\mu + 2\right]\left[\frac{z}{z_0} - 1\right]^2 \quad [85]$$

Thus the solution can be represented by a one-parameter family of curves, and these are shown plotted in Figs. 7 and 8.

VII Conclusions

A general approximate mathematical method for solving heat-transfer problems utilizing the heat-balance integral has been presented and applied to five problems involving a change of phase. By representing the temperature distribution by a quadratic, the results have been expressed in closed analytical form. Comparisons with known solutions have been made whenever possible, and ultimately, all the solutions have been presented in graphical form.

The technique presented in this paper is applicable to a wide variety of heat-transfer problems, but finds its widest application to those which are nonlinear, and must, therefore, be solved either numerically or approximately. There are many such problems which do not warrant the tedious labor involved in numerical integration of the heat-conduction equation, and it is here that the method of the heat-balance integral will find its widest acceptance.

Bibliography

1 "Conduction of Heat in Solids," by H. S. Carslaw and J. C. Jaeger, Oxford University Press, London, England, first edition, 1947, pp. 56–57, 71–74.

2 "Stefan-Like Problems," by G. W. Evans, E. Isaacson, and J. K. L. MacDonald, *Quarterly of Applied Mathematics*, vol. 8, 1950, pp. 312–319.

3 "Heat Conduction in a Melting Solid," by H. G. Landau, *Quarterly of Applied Mathematics*, vol. 8, 1950, pp. 81–94.

4 "Finite Difference Approach to Some Heat Conduction Problems Involving Changes of State," by C. A. Forster, English Electric Company, Ltd., Report L A.t. 059, April 6, 1954.

5 "An Expansion Method for Parabolic Partial Differential Equations," by J. W. Green, National Bureau of Standards, *Journal of Research*, vol. 51, September, 1953, pp. 127–132.

6 "New Methods in Heat Flow Analyses With Application to Flight Structures," by M. A. Biot, IAS Preprint No. 661, 1957.

7 "Zur naherungsweisen Integration der Differentialgleichunger der laminaren Grenzschicht," by K. Pohlhausen, *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 1, 1921, pp. 252–258.

8 "Boundary Layer Theory," by H. Schlichting, McGraw-Hill Book Company, Inc., New York, N. Y., 1955, Chapter 12.

VIII APPENDIX

Melting With Fixed Boundary Temperature, Cubic Profile. Assuming a cubic temperature profile, and applying Conditions [16], [19], and [24], there is obtained

$$\frac{u}{\kappa A} = -q\frac{(x-s)}{s} + \frac{1}{2}q^2\frac{(x-s)^2}{s^2} - \frac{1}{2}(\mu - 2q - q^2)\frac{(x-s)^3}{s^3} \quad [86]$$

where q is an unknown positive function of time to be determined. Substituting into Equations [20] and [21] the heat-balance integral becomes

$$s\frac{d}{dt}\{s[q^2 + 6q + 3\mu + 24]\} = -12\kappa[q^2 + 4q - 3\mu] \quad [87]$$

The condition at the melt line, Equation [18], becomes

$$s\frac{ds}{dt} = \kappa q \quad [88]$$

The solution is q = const; then eliminating $s(ds/dt)$ between Equations [87] and [88] there is obtained a functional relationship between q and μ

$$\mu = \frac{q^3 + 18q^2 + 72q}{3(12 - q)} \quad [89]$$

From Equation [88], the solution for s is of the form of Equation [29] where

$$\frac{K}{2\sqrt{\kappa}} = \sqrt{\frac{q}{2}} \quad [90]$$

Equations [89] and [90] represent the solution for $K/2\sqrt{\kappa}$ in terms of μ in parametric form. The result is shown plotted in Fig. 1.

The Biotechnical Problem of the Human Body as a Heat Exchanger

By L. P. HERRINGTON,[1] NEW HAVEN, CONN.

The physical and engineering properties of inanimate objects as heat exchangers have been the subject of long study. Within the past 25 years, many factors have brought biological disciplines into intimate contact with the formally similar problem of heat exchange between a living body and its environment. A Committee on Biotechnology of the Heat Transfer Division of this Society has been organized recently with the intention of advancing and standardizing useful engineering descriptions of biological heat-exchange problems. Such problems presently complicate engineering design in which the human link is a critical element in total function of man and machine. The paper demonstrates that a large body of calorimetric data on the human heat exchanger can be summarized in statistically derived empirical equations. These equations obviate the need for special physiological knowledge required of the engineer who would make such computations from the classical equations of heat loss.

Introduction

As a biologically oriented member (1)[2] of the Committee on Biotechnology of the Heat Transfer Division, the author is convinced that a useful initial step toward an understanding of biological heat-exchange problems would be to provide the engineer with convenient equational condensations of human calorimetric data (2, 3, 4, 5).

In these and other calorimetric studies reliable determinations have been made of radiation, convection, and evaporative coefficients of the human heat exchanger. Such coefficients have been applied to the classical heat-exchange equations with which engineers are thoroughly familiar. In this sense, the analogy between the animate and inanimate heat-exchange studies is exact.

Human Organism as a Heat Exchanger

However, the human organism differs from other heat exchangers in that its heat-regulation reflexes (2) may be said to be a servomechanism with a complicated yet consistent method of responding to heat or cold stress. The nature of this patterned response is such that in contrast to inanimate heat exchangers, the following properties of the human heat exchanger may undergo complex interrelated alterations over even a small range of ambient temperatures (area 70–90 F):

(*a*) Heat input to the exchanger.

(*b*) Alteration of the conductance of the peripheral material of the exchanger.

(*c*) Alteration of the total conductance from surface to ambient surround.

(*d*) Shift of heat-dissipating load from a low surface-volume ratio exchanger segment (the trunk) to a segment of higher relative surface (extremities).

(*e*) Conversion of the surface process of the exchanger from a dry-heat transfer to a combined dry-heat and evaporative process.

Experimentally Derived Coefficients for Conventional Equations of Heat Loss as Applied to the Human Exchanger

It has been absolutely necessary to make calorimetric studies of the human heat exchanger by use of the classical heat relations. It must be obvious, however, from an inspection of the servoproperties just listed that a very large amount of physiological study is required if an engineer desires to apply the classical relations to an immediate problem. Briefly, this is because he must be able to decide what the control state of the human heat servomechanism is for a particular ambient condition, or a particular state of work stress in the organism.

This application difficulty may be realized easily by inspecting the following classical heat-exchange relations for the human body. In making this inspection the question may be asked as to how one proceeds with the computation, taking due account of the features of heat-exchanger behavior noted (*a–e*) as continuously varying properties of the experimental object.

Radiation Exchange. The equation for heat transfer by radiation (2) between the unclothed human body and the environment is given by

$$H_R = 1.37 \times 10^{-11}(T_s^4 - T_e^4)tAf\epsilon, \text{ kg cal/hr} \ldots\ldots [1]$$

where T_s = average skin temperature (deg C + 273), T_e = average radiant environmental temperature (+ 273), t = seconds in one hour, A = DuBois surface area, f = ratio of effective radiating surface to the DuBois surface area (0.78 for the unclothed adult lying in anatomical position), and ϵ = emissivity of the environment.

Conduction Exchange. Clothing and other factors normally reduce human conductive heat to a small fraction of the total exchange. However, the familiar classical equation is frequently applied (6, 7) to the problem of computing the alteration in the conduction of heat from the interior of the body to the surface. (See item *b* in the list of body heat-exchanger properties altered by biological servocontrol mechanisms.)

Forced Convection. Some degree of forced convection is generally present in human heat exchange. The arrangement of variables which best fits with the experimental facts and theoretical considerations has been analyzed by the author and his colleague with reference to biological heat exchange (2). The equational expression used is

$$H_c = \frac{K}{D}\left(1 + \alpha\left(\frac{DV\rho}{\mu}\right)^{1/3} + b\left(\frac{DV\rho}{\mu}\right)\right)\Delta Tt \ldots\ldots [2]$$

where H_c = heat loss by convection, D = characteristic dimension of object (for example, the diameter of a sphere or a cylinder), V = velocity of the gas, μ = viscosity—a factor concerned in the mobility of the gas molecule, ρ = density, K = thermal conductivity, T = temperature difference between

---

[1] The John B. Pierce Foundation.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 23, 1956. Paper No. 57—SA-5.

the warm surface and the air, $\Delta T = T_s - T_A$, and t = time.

Terms α and b are constants depending upon the particular units used. It is convenient to reduce all surfaces to equivalent cylinders or spheres, since most of the experimental work has been done by engineers who are interested in convection losses from pipes. Neglecting everything but convection, the adult human body loses heat like a cylinder 7 cm (c. 3 in.) in diameter ($\alpha = 0.407$, $b = 0.00123$, if velocity is expressed in miles per hour, diameter in inches, and convective heat loss in kg cal/m²/hr/deg F) or like a sphere 15 cm in diameter (8). The study of convection losses from the formulas developed by the engineers can be made most easily by considering the body as a 3-in. cylinder or a 6-in. sphere.

TABLE 1 MEAN[a] VALUES OF CALORIMETRIC DATA, MATURE MALE SUBJECTS NORMAL CLOTHING, SEATED POSTURE

	Skin temp.	Air temp.	Radiant temp.	Operative temp.[a]	Metabolism	Evaporation
Mean	85.25	49.02	71.52	61.61	91.55	−20.86
Std deviation	3.26	10.70	16.15	10.17	7.72	5.70
Population[b]	180	180	180	180	180	180

[a] Operative temperature resembles in principle the process of reducing a gas to a reference volume and pressure. It may be understood in a sensory sense as the temperature of an enclosure with walls and air at the same temperature, and with an air movement of 15 to 20 fpm, with relative humidity of 50 per cent. Technically it is given by the equation

$$T_0 = \frac{k_R T_W + k_C T_A}{k_0} \quad [4]$$

where T_0 = operative temperature, T_W = radiant or wall temperature, k_R and k_C = subject radiation and convection heat-exchange constants as determined by calorimeter experiments, and $k_0 = k_R + k_C$.

[b] Each item in the population is an exposure to a given calorimeter combination of air and radiant temperature for 3 hr.

NOTE: Temperatures, deg F; metabolism and evaporation, kg cal per hr per man; avg height 180 cm; avg weight 70.8 kg; avg DuBois surface area 1.76 sq meters; avg radiation area 1.34 sq meters.

Evaporation. The rate and amount of evaporative heat exchange is a variable of wide range in the biological heat-exchange processes. Knowledge of the physiological stress response for particular environments is an essential prerequisite for this calculation (2, 3, 9, 10). The physical elements of the process are given by

$$H_v = (\omega\mu)A(E_s - \mathrm{RH}\ E_A) \text{ kg cal/hr}$$

or

$$(\omega\mu) = \frac{H_v}{A(E_s - \mathrm{RH}\ E_A)} \quad [3]$$

where ω = fraction of body area that is completely wet, μ = proportionality factor containing the vaporization constant and the factors which depend on air velocity and direction, H_v = heat loss by vaporization, A = total body area, and RH = relative humidity.

Engineering Utility of Human Heat-Transfer Equations Which Include by Statistical Means Changes in Physical Properties of Exchanger

The biological reference tools needed to apply Equations [1], [2], and [3] to concrete situations are clearly very extensive. There is thus an excellent argument for computing multiple regression equations which apply to specific ranges of human adjustment to heat stress. As a result of the approximate linearity (11) of the interrelations of important physiological and environmental variables within regions of ambient stress which stimulate a typical and progressive pattern of alteration in the biological heat-exchanger properties (crudely, conditions felt as cold, neutral, or warm in a sensory sense), it is possible to condense calorimetric data into very compact expressions which combine these effects. Such methods should provide the most direct summary of the effects due to the operation of the classical physical laws on a heat exchanger reflexly equipped to maximize or minimize the integral heat-transfer effect attributable to each of the avenues of physical heat loss.

Derivation of Linear First-Order Equations in Five Variables Relating Ambient Air and Radiant Temperatures to Skin Temperature, Metabolic Heat Input, and Evaporative Heat Loss

The most frequent engineering requirement involving human heat loss requires the estimation of a resultant skin temperature re the problem of heat tolerance. Such design situations are generally too unusual to be settled by reference to data for the circumstances obtaining in ordinary air-conditioning problems. In other instances the problem is to estimate the stress effect of an increase in the radiant temperature of an environment, or to estimate the highest level of activity (heat input) consistent with fixed environmental heat effects.

In all such instances, a linear multivariable equation, permitting values to be fixed for four of the chief variables with solution in terms of a remaining variable, is of great usefulness.

The calorimetric log of human heat-exposure experiments of the author and associates at the Pierce Foundation Laboratory of Hygiene has been abstracted for the basic data of such a computation. (See reference 3 and references therein). In 180 calorimeter experiments on mature, normally clothed male subjects, the mean values of Table 1 were found for the group of 3-hr exposures. All of these experiments were under conditions which do not stimulate positive sweating.

The mean values given in Table 1 represent 180 numerical values distributed among six variables. To express these 180 values in a single equation requires the determination of the regressions between every possible combination of the six variables. Since operative temperature is derived from air and radiant temperature, the calculation program was reduced to one dealing with the five remaining variables. The Pearson product-moment method was applied to determine the 10 least-square solutions existing in the system of variables. From these intercorrelations the four partial regressions were determined, representing the relation between four pairs of variables with the influence of the remaining three variables removed mathematically.

This is a tedious operation and time-consuming, but the end result is very efficient in that it enables us to derive a five-element equation from which the entire table of 180 data entries may be regenerated with surprisingly small deviation between the values of the regenerated table and the actual observational data from the calorimeter.

An Equation in Five Variables Representing Interrelations Existing Between Mean Skin-Surface Temperature of Human Heat Exchanger and Important Ambient and Internal Factors

The final result of the mathematical analysis of the distribution of calorimeter data associated with mean values of Table 1 is

$$\underset{(T_S)}{X_1} = 0.286\underset{(T_A)}{X_2} + 0.142\underset{(T_W)}{X_3} + 0.105\underset{(M)}{X_4} + 0.092\underset{(E)}{X_5} + 53.39 \quad [5]$$

where T_S, T_A, T_W, refer, respectively, to mean skin surface, air, and wall temperatures, and M and E to metabolism and evaporation. In solutions of the equation the units of the variables as given in Table 1 must be used.

The high efficiency of the equation may be judged from the fact that the multiple regression coefficient for the equation is

$$R_{1.2345} = 0.908 \pm 0.03$$

where R designates the multiple correlation coefficient.

We may check the accuracy of the equation by substituting in Equation [5] the mean values of Table 1.

TABLE 2

Variable	Mean value	Coefficient Equation [5]	Product	Summation
T_A.....	49.02	0.286	+14.019	14.019
T_W.....	71.52	0.142	+10.156	24.175
M.....	91.55	0.105	+ 9.613	33.788
E.....	−20.86	0.092	−1.919	31.869
Constant	...	...	+53.39	....
T_S.....	*85.25			*85.259

From Table 2 it may be seen that the mean surface temperature of the human heater exchanger as an equational function of the two ambient and two physiological variables is identical to four significant figures with the grand mean of the 540 hr of observation. Since the multiple correlation coefficient of the solution is 0.908 ± 0.03 it serves no purpose to reproduce here the full table comparing 180 environmental and computed values. The agreement is excellent through the experimental range of values indicated by the means and standard deviations of Table 1.

Application of Equation [5]

Equation [5] may be applied accurately to any seated-activity situation in which the crude average of air and radiant temperatures is between 50 and 80 F and with air movements of the order of 10 to 20 fpm, with occupants wearing normal male attire, or approximately seven pounds of clothing. In the absence of separate measurements of air and radiant temperature, or in the presence of air movements up to 100 fpm, it may be applied with approximate accuracy. In this latter case the reading of a black globe thermometer should be used to estimate T_0 (the combined T_A, T_W) effect, and the reading of this instrument substituted in both the T_A and the T_W terms of the equation.

In computations for other work levels, with metabolism higher or lower than 90 kg cal per hr per 1.76 of body surface, the E-factor for the new M heat input should be estimated as proportional to −20.86/91.55 or −0.228 of the new heat input figure—provided the trial solution of Equation [5] does not yield a skin-surface temperature above 94 F. Between 94 and 95 F, as a tissue surface-boundary condition between body and environment, a rapidly progressive increase in positive evaporative regulation occurs. In addition, large increments occur in the convective effect of peripherally flowing blood. As a result, a separate equation must be used to describe the interaction of the variables of Table 1 above a skin temperature of 94 F.

Such an equation is now being derived by the author. In addition, equations similar to Equation [5] (based upon the same data) are nearing solution, in which the predicted value is the surface temperature of the head, the upper extremities, the trunk, and lower extremities.

The local segmental temperature of the legs is important since they are effective convectors, and as their temperature increases rapidly with rising temperatures, the point at which the leg surfaces approach 95 F is a valuable index of the inflection of properties in the total human heat exchanger. Since leg surface temperature is easy to measure, users of Equation [5] who may wish to enter an equation with a measured value of skin temperature (and solve for the required ambient factor) may do so if a similar solution is available in terms of the surface temperature of the extremities

By methods similar to those described with reference to Equation [5] such an equation has been derived. It is

$$\left.\begin{array}{l} X_6 = 0.595X_7 - 0.133X_8 + 0.206X_9 + 0.245X_5 + 33.61 \\ \quad\searrow\quad\quad\searrow\quad\quad\searrow\quad\quad\searrow\quad\quad\searrow \\ \quad(T_L)\quad\quad(T_0)\quad\quad(T_W - T_A)\;(M)\quad\quad(E) \end{array}\right\} \quad [6]$$

where T_0 is the black-globe thermometer temperature, and $(T_W - T_A)$ is the difference, if any, by which radiant temperature exceeds air temperature.

Summary

Engineering-design problems now frequently involve a man-machine problem in which design is affected by human tolerance to cold or heat stress. Special physical features of these environments (12) frequently render the conventional data of thermal-engineering rules inapplicable. It has been shown in this paper that a large body of calorimetric data on the human heat exchanger can be summarized in statistically derived empirical equations in five variables. Such equations when applied within the ranges indicated greatly reduce the labor of computation, and to a large degree obviate the need for special physiological knowledge required of the engineer who would make such computations from the classical equations of heat loss.

This analysis is presented by the author to serve the interests of the Society's Committee on Biotechnology, and will be followed by similar work designed to provide a set of five variable equations covering all important variations of human work load and ambient heat conditions which may affect engineering design through human restrictions imposed by thermal stress in man-machine systems.

BIBLIOGRAPHY

1 "Biophysical Adaptations of Man Under Climatic Stress. Recent Studies in Bioclimatology," by L. P. Herrington, *Meteorological Monographs*, vol. 2, no. 8, 1954, pp. 30–42.

2 "Temperature and Humidity in Relation to the Thermal Interchange Between the Human Body and the Environment," by L. P. Herrington and J. D. Hardy, Chapter 13, "Human Factors in Undersea Warfare," National Research Council, Washington, D. C., 1949.

3 "Temperature and Human Life," by C. E.-A. Winslow and L. P. Herrington, Princeton University Press, Princeton, N. J., 1949, 266 pp.

4 "Basic Procedures in the Calculation of the Heat Exchange of the Clothed Human Body," by L. P. Herrington, *Yale Journal of Biology and Medicine*, vol. 19, 1947, pp. 735–755.

5 "Physiology of Heat Regulation and the Science of Clothing," by W. B. Saunders, Symposium Volume, edited by L. H. Newburgh, auspices of the Medical Division, N.R.C., 1949.

6 "The Relative Influence of Radiation and Convection Upon Vasomotor Temperature Regulation," by L. P. Herrington, C. E. A. Winslow, and A. P. Gagge, *American Journal of Physiology*, vol. 120, 1937, pp. 133–143.

7 "The Heat Regulation of Small Laboratory Animals at Various Environmental Temperatures," by L. P. Herrington, *American Journal of Physiology*, vol. 129, 1940, pp. 123–139.

8 "Publication of the Climatology and Environmental Protection Branch," by J. H. Plumer, Office of the Quartermaster General, Washington, D. C., August 25, 1944.

9 "Bedeutung und Messung der Oberflächenfeuchte für Transpirationsanalyse," by K. Büttner, *Biologischen Zentralblatt*, vol. 55, 1935, p. 356.

10 "A New Physiological Variable Associated With Sensible and Insensible Perspiration," by A. P. Gagge, *American Journal of Physiology*, vol. 120, 1937, pp. 277–286.

11 "The Linearity Criterion as Applied to Partitional Calorimetry," by A. P. Gagge, *American Journal of Physiology*, vol. 116, 1936, pp. 656–668.

12 "The Physiological Engineering of Human Habitation," by L. P. Herrington, Proceedings of the Tenth Annual Builders Conference, Department of Architecture, University of Illinois, Urbana, Ill., January 13, 1955.

Discussion

E. F. Adolph.[3] Heat gains and losses by clothed men are represented in the author's equation. The device of using linear approximations for the additive functions involved in the equation is a familiar and effective one. Anyone can check the equation itself for accuracy only by having access to data similar to those used by the author in deriving it. The equation also serves to list factors which must be taken into account in any air-conditioning problem, and conveys some insight into the general concepts of man's heat exchanges. Use of the particular equation given must be restricted to men under the conditions specified, which are seated at rest, wearing average woolen clothing, in the range of 50 to 80 F air and wall temperatures, with air motion of 10 to 20 fpm. In this restricted range the equation will be a practical aid to engineering computations.

A. H. Woodcock.[4] The author has chosen to treat the problem of man as a heat exchanger from an empirical viewpoint. The writer will comment on this method. An empirical relationship or equation is one that is fitted in the best and simplest manner to the known data. It is not at all necessary that it should have the true form of the relationship. As such it is useful for interpolation, but should be used with extreme caution for extrapolation.

As an example, a series of experimental points might be described empirically by the equation of a straight line with considerable accuracy although they actually lie on a hyperbolic curve at some distance from its focus.

Interpolation with the empirical linear relationship would give accurate results, but extrapolation to points near the focus obviously would result in completely erroneous predictions.

This paper is an excellent example of a well-organized presentation of an empirical relationship.

The author has been extremely careful to state quite clearly the limits over which Equations [5] and [6], the derived empirical relations, apply. These equations are straightforward and easy to apply, and are therefore of considerable practical use. If one were to develop the true equations from those which involve radiation, conduction, convection, and evaporation, a much more complicated and unwieldy relationship would result.

The advantages of the empirical method depend, of course, on the skill of the person setting up the relationship. In this paper the author has demonstrated that skill. He has eliminated the need for determining all the biological implications which occur in the problem and in which the reader may not be interested. In its place he has given a simple straightforward equation and has outlined the conditions under which it may be used.

Author's Closure

The discussions describe correctly the properties of this method of analysis and the area of application for equations of the type developed in this study.

[3] Department of Physiology, University of Rochester, School of Medicine and Dentistry, Rochester, N. Y.

[4] Quartermaster, Research and Development Command, Natick, Mass.

Transient Free Convection From a Vertical Flat Plate

By ROBERT SIEGEL,[1] CLEVELAND, OHIO

The method of characteristics is employed to obtain solutions to the time dependent free-convection equations of momentum and energy placed in integral form (Karman-Pohlhausen method). Two boundary conditions are considered for a vertical flat plate of infinite width and semi-infinite length which is initially at ambient temperature in quiescent fluid: (*a*) The plate is suddenly raised to a uniform higher temperature, and (*b*) the plate suddenly begins to produce a uniform heat flux at its surface. The results yield the time required for steady flow to be established as a function of position along the plate. Heat-transfer coefficients are obtained for the initial stage of motion during which the convective process is one dimensional. The approximate velocity and temperature profiles obtained from the analysis are compared with more precise solutions of the differential equations for the initial stage of motion and for steady state.

Nomenclature

The following nomenclature is used in the paper:

c_p = specific heat at constant pressure, Btu/(lb deg F)
g = acceleration of gravity, ft/sq sec
h = local coefficient of heat transfer, Btu/(sec, sq ft, deg F)
k = thermal conductivity, Btu/(sec, ft, deg F)
L = arbitrary characteristic length, ft
q = heat flux per unit area at plate surface, Btu/(sec, sq ft)
t = temperature, deg F
U = velocity in X-direction, fps
U_1 = characteristic velocity in boundary layer defined by

$$U = U_1 f\left(\frac{Y}{\Delta}\right), \text{ fps}$$

u = dimensionless velocity defined as UL/α
u_1 = dimensionless velocity defined as U_1L/α
V = velocity in the Y-direction, fps
X = co-ordinate of height along flat plate measured upward from lower edge, ft
x = dimensionless co-ordinate defined as X/L
Y = co-ordinate normal to plate and measured from plate surface, ft
α = thermal diffusivity defined as $k/\rho c_p$, sq ft/sec
β = volumetric coefficient of expansion, deg F^{-1}
Δ = boundary-layer thickness, ft
δ = dimensionless boundary-layer thickness defined as Δ/L
ζ = a function of η defined by Equation [83]
η = parameter defined as $Y/2\,(\alpha T)^{1/2}$
θ = temperature at an arbitrary point in boundary layer minus ambient temperature, deg F
θ_w = wall temperature minus ambient temperature, deg F
μ = coefficient of viscosity, lb/(sec ft)
ν = kinematic viscosity, sq ft/sec
ρ = density, pcf
T = time, sec
τ = dimensionless time defined as $T\alpha/L^2$
Φ = dimensionless temperature ratio, θ/θ_w
χ = parameter defined as $(Y/X)\,(\mathrm{Gr}_X/4)^{1/4}$
χ^* = parameter defined as $(Y/X)\,(\mathrm{Gr}_X{}^*/5)^{1/4}$

Subscripts

∞ = ambient condition outside of boundary layer
w = location at surface of heated plate

Dimensionless Groups

Gr_X = local Grashof number, $g\beta\theta_w X^3/\nu^2$
$\mathrm{Gr}_X{}^*$ = modified local Grashof number based on q, $g\beta q X^4/k\nu^2$
Nu_X = local Nusselt number, hX/k
Pr = Prandtl number, $c_p\mu/k$
Ra_L = Rayleigh number based on L, $g\beta\theta_w L^3 c_p\mu/\nu^2 k$
$\mathrm{Ra}_L{}^*$ = modified Rayleigh number based on q and L, $g\beta q L^4 c_p\mu/\nu^2 k^2$

Special Symbols

I_+, I_- = designate the two families of characteristic lines
$\mathrm{I}_+{}^\circ$, $\mathrm{I}_-{}^\circ$ = two characteristic lines which pass through origin of the τ-x plane
II_+, II_- = designate the two characteristic equations

Introduction

In the design of nuclear-reactor fuel elements it is necessary to consider their temperature behavior during various types of power transients. Under some conditions, for example during a coolant-pump failure, the heat removal to the cooling fluid may be solely by free convection. To obtain some insight into the convective process it is necessary to consider how the free-convection boundary layer is influenced when the heated surface is undergoing a thermal transient.

One phase of this problem was considered by Illingworth (1)[2] who studied one-dimensional free convection about an infinite plate undergoing a step-function change in temperature. For a plate of semi-infinite length, Sugawara and Michiyoshi (2) treated a step-function change in wall temperature by using a method of successive approximations in which the heat transfer was taken to be purely by conduction for the first approximation. The velocity and temperature distributions from this solution were then utilized in the differential equations to obtain a second approximation. The corrections resulting from the second approximation are only given for very short times, and it is not known if the second approximation will converge sufficiently well to the steady-state solution at large times.

The method of successive approximations in (2) is similar to that employed by Blasius for determining the boundary-layer growth on a body started impulsively from rest.[3] Schuh (4)

[1] Aeronautical Research Scientist, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics. Assoc. Mem. ASME.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 22, 1957. Paper No. 57—SA-8.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

[3] Reference (3) p. 181.

investigated the same type of problem as Blasius by using an integral-momentum (Karman-Pohlhausen) method and obtained solutions by the method of characteristics. This type of approach was utilized in the present paper to study transient free convection on a vertical flat plate of infinite width and semi-infinite length. Two boundary conditions are considered for a plate initially at ambient temperature in quiescent fluid: (*a*) The plate is suddenly raised to a higher uniform temperature, and (*b*) the plate suddenly begins to generate a uniform heat flux at its surface.

The equations of momentum and energy are integrated by assuming approximate velocity and temperature profiles. This results in two simultaneous partial differential equations which are to be solved for the boundary-layer thickness δ and the velocity u_1 as a function of time and position along the plate. The equations are shown to be hyperbolic in type and a solution is obtained by the method of characteristics. In the early stages of motion, heat is transferred by conduction only, and a one-dimensional diffusion-type solution results. The characteristic lines through the origin of the τ-x plane show how the influence of the plate leading edge propagates into the fluid and causes the one-dimensional motion to adjust to the steady-state condition.

In the analysis the case of the constant-temperature plate is presented in some detail for a particular choice of approximate velocity and temperature profiles. To determine the sensitivity of the method to the profile shapes, a second choice of profiles was made and a summary of results is given. The equations are then rearranged to accommodate a constant heat-flux boundary condition and results are obtained for two sets of profile shapes.

Plate at Uniform Temperature

Basic Equations. The first situation considered is the flow development about a vertical flat plate which is initially at thermal equilibrium in quiescent fluid, and is then raised in a step function fashion to a uniform temperature, t_w. The X-direction extends vertically upward from the lower edge of the plate, while Y is measured normal to the plate and away from the plate surface. The time-dependent free-convection equations of momentum and energy in integral form are

$$\frac{\partial}{\partial T}\int_0^\Delta U dY + \frac{\partial}{\partial X}\int_0^\Delta U^2 dY = g\beta\int_0^\Delta (t - t_\infty)dY - \nu\left.\frac{\partial U}{\partial Y}\right|_{Y=0} \quad \ldots\ldots [1]$$

$$\frac{\partial}{\partial T}\int_0^\Delta (t - t_\infty)dY + \frac{\partial}{\partial X}\int_0^\Delta (t - t_\infty)UdY = -\alpha\left.\frac{\partial t}{\partial Y}\right|_{Y=0} \quad \ldots\ldots [2]$$

Except for the terms involving time derivatives, Equations [1] and [2] are the same as the equations given by Eckert.[4] As in (5) it has been assumed that the fluid properties are constant except for variable density in formulating the buoyancy term, and that the thermal and hydrodynamic boundary layers are equal in thickness.

To solve the set of equations, approximate velocity and temperature distributions are assumed, and these are taken to be reasonably good approximations throughout the entire transient-flow development. This last approximation can be checked to some extent by comparisons with exact solutions given later in the paper. Following Eckert,[5] we let the velocity and temperature profiles be

$$\frac{U}{U_1} = \frac{Y}{\Delta}\left(1 - \frac{Y}{\Delta}\right)^2 \quad \ldots\ldots [3]$$

$$\frac{t - t_\infty}{t_w - t_\infty} = \frac{\theta}{\theta_w} = \left(1 - \frac{Y}{\Delta}\right)^2 \quad \ldots\ldots [4]$$

When these profiles are inserted into Equations [1] and [2] and the integrations performed, the results are (in dimensionless form)

$$\frac{1}{12\mathrm{Pr}}\frac{\partial}{\partial\tau}(u_1\delta) + \frac{1}{105\mathrm{Pr}}\frac{\partial}{\partial x}(\delta u_1^2) = \frac{1}{3}\mathrm{Ra}_L\,\delta - \frac{u_1}{\delta} \quad \ldots\ldots [5]$$

$$\frac{\partial\delta}{\partial\tau} + \frac{1}{10}\frac{\partial}{\partial x}(u_1\delta) = \frac{6}{\delta} \quad \ldots\ldots [6]$$

where Ra_L is the Rayleigh number based on the length L.

These two equations are to be solved simultaneously for the velocity u_1 and the boundary-layer thickness δ as functions of τ and x. A method of solution for a set of this type is given in (6), section 22, and the details of application to the present problem are given in Appendix 1. As shown in the Appendix, Equations [5] and [6] are hyperbolic in type and hence a solution can be found by utilizing the method of characteristics. The slopes of the characteristic lines are found to be proportional to the velocity u_1. Letting I_+ and I_- designate the two families of characteristics, we have

$$\mathrm{I}_+ \quad \frac{dx}{d\tau} = 0.155\,u_1 \quad \ldots\ldots [7]$$

$$\mathrm{I}_- \quad \frac{dx}{d\tau} = 0.0739\,u_1 \quad \ldots\ldots [8]$$

The characteristic equations which apply along the characteristic lines are designated by II_+ and II_-

$$\mathrm{II}_+ \quad \delta\frac{du_1}{d\tau} + 0.261u_1\frac{d\delta}{d\tau} + (4.43 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\mathrm{Pr}\delta = 0 \quad \ldots\ldots [9]$$

$$\mathrm{II}_- \quad \delta\frac{du_1}{d\tau} - 0.547u_1\frac{d\delta}{d\tau} + (9.28 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\mathrm{Pr}\delta = 0 \quad \ldots\ldots [10]$$

Solutions of Characteristic Equations. When a vertical plate is suddenly raised to a uniform temperature, the heat transfer to the surrounding fluid is initially by pure conduction and hence is equivalent to the heat conducted into a semi-infinite solid when its surface temperature is suddenly increased. This situation arises from the fact that the fluid sufficiently far from the leading edge behaves as if the plate were infinite in length, so that the velocity distribution in this region is independent of x and hence the convective heat transfer is zero. The two-dimensional influence which causes the boundary-layer growth to vary with x gradually propagates from the leading edge and begins to alter the one-dimensional flow configuration at a different time for each position along the plate. Thus for each x it would be expected that, for sufficiently small times, the solution would be dependent on time only. On this basis a trial solution of the form

$$\delta = C_1\tau^m \quad \ldots\ldots [11]$$

$$u_1 = C_2\tau^n \quad \ldots\ldots [12]$$

was substituted into Equations [9] and [10]. For the resulting equations to be independent of τ it was found that $m = 1/2$ and

[4] Reference (5), p. 160

[5] Ibid., p. 159.

$n = 1$. The equations were then solved simultaneously for C_1 and C_2 giving the results

$$\delta = \sqrt{(12)}\,\tau^{1/2} \quad \text{or} \quad \Delta = \sqrt{(12)}\,(\alpha T)^{1/2} \ldots\ldots [13]$$

$$u_1 = \frac{4\mathrm{Ra}_L\mathrm{Pr}}{1.5 + \mathrm{Pr}}\,\tau \quad \text{or} \quad U_1 = \frac{4g\beta\theta_w}{1.5 + \mathrm{Pr}}\,T \ldots\ldots [14]$$

These results also can be obtained directly from the original Equations [5] and [6] if the derivatives with respect to x are set equal to zero. The characteristic solution is necessary, however, to determine the domain of the τ-x plane in which the solution applies.

At sufficiently large times, the characteristic equations also should yield a steady-state solution in which u_1 and δ are functions of x only. To obtain this result Equations [9] and [10] are first rewritten in the form

$$\mathrm{II}_+ \quad \delta\frac{du_1}{dx}\frac{dx}{d\tau} + 0.261\,u_1\frac{d\delta}{dx}\frac{dx}{d\tau} + (4.43 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\,\mathrm{Pr}\delta = 0 \ldots\ldots [15]$$

$$\mathrm{II}_- \quad \delta\frac{du_1}{dx}\frac{dx}{d\tau} - 0.547\,u_1\frac{d\delta}{dx}\frac{dx}{d\tau} + (9.28 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\mathrm{Pr}\delta = 0 \ldots\ldots [16]$$

Then $dx/d\tau$ is eliminated from Equation [15] by substituting Equation [7], and from Equation [16] by use of Equation [8], and a solution is tried of the form

$$\delta = C_3x^r$$

$$u_1 = C_4x^s$$

For Equations [15] and [16] to be satisfied the exponents must be $r = 1/4$ and $s = 1/2$. The equations are then solved for C_3 and C_4 with the results

$$\delta = 3.93\,(0.952 + \mathrm{Pr})^{1/4}\,(\mathrm{Ra}_L\mathrm{Pr})^{-1/4}\,x^{1/4} \ldots\ldots [17]$$

$$u_1 = 5.17\,(0.952 + \mathrm{Pr})^{-1/2}\,(\mathrm{Ra}_L\mathrm{Pr})^{1/2}\,x^{1/2} \ldots\ldots [18]$$

This is in agreement with the steady-state solution for free convection on a vertical flat plate at uniform temperature as given in (5).[6] The characteristics will be used to determine the domain of the τ-x plane in which this solution applies.

Regions of the τ-x Plane in Which Solutions Apply. Consider the two characteristic lines which pass through the origin of the τ-x plane as shown in Fig. 1. From Equations [7] and [8] it is noted that the I_+ lines in the τ-x plane have smaller slopes than the I_- lines. Hence the two lines passing through the origin (designated by $\mathrm{I}_+°$ and $\mathrm{I}_-°$) divide the plane into three regions.

In region A, along and below the $\mathrm{I}_+°$ line, values of u_1 and δ can be obtained at the intersections of sets of characteristics which originate on the x-axis. The solutions of the characteristic equations must satisfy the boundary conditions that u_1 and δ are zero for all x along the x-axis ($\tau = 0$). These requirements are satisfied by the solution dependent on time only which is given in Equations [13] and [14]. Hence in region A the heat transfer is by purely one-dimensional conduction, and the $\mathrm{I}_+°$ line gives the time, corresponding to each position along the plate, at which the one-dimensional process ends and the effect of the leading edge begins to influence the velocity and temperature distributions.

The equation of the $\mathrm{I}_+°$ line can be found by integrating Equation [7] with u_1 obtained from Equation [14], and with the condition that $x = 0$ at $\tau = 0$.

[6] Reference (5), p. 161.

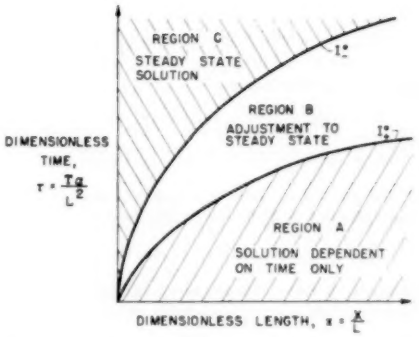


FIG. 1 REGIONS ON THE τ-x PLANE IN WHICH PARTICULAR SOLUTIONS APPLY

After rearrangement this yields

$$\tau = 1.80\,(1.5 + \mathrm{Pr})^{1/2}\,(\mathrm{Ra}_L\mathrm{Pr})^{-1/2}\,x^{1/2}$$

or

$$T = 1.80\,(1.5 + \mathrm{Pr})^{1/2}\,(g\beta\theta_w)^{-1/2}\,X^{1/2} \ldots\ldots [19]$$

Equation [19] then gives the time at which purely one-dimensional heat conduction is terminated for each position along the plate.

In region C along and above the $\mathrm{I}_-°$ line, values of u_1 and δ are found at intersections of characteristic lines originating along the τ-axis. The solutions of the characteristic equations must satisfy the boundary conditions that for any τ, at $x = 0$ (the leading edge of the plate), the values of u_1 and δ must be zero. These conditions are satisfied by the steady-state solution, and thus the equation of the $\mathrm{I}_-°$ line gives the time at which steady state is reached as a function of position along the plate.

The equation of $\mathrm{I}_-°$ is found by integrating Equation [8] with u_1 given by the steady-state solution, Equation [18], and subject to the condition $x = 0$ at $\tau = 0$. This yields after rearrangement

$$\tau = 5.24\,(0.952 + \mathrm{Pr})^{1/2}\,(\mathrm{Ra}_L\mathrm{Pr})^{-1/2}\,x^{1/2}$$

or

$$T = 5.24\,(0.952 + \mathrm{Pr})^{1/2}\,(g\beta\theta_w)^{-1/2}\,X^{1/2} \ldots\ldots [20]$$

which is the time required to reach steady state.

As a numerical example we shall calculate the time required to establish the flow pattern in air over the first foot of a plate which has been suddenly raised from 70 to 270 F. Then

$$T = 5.24\,(0.952 + 0.70)^{1/2}\left[\frac{530}{32.2\,(200)}\right]^{1/2}(1)^{1/2} = 1.93 \text{ sec}$$

Alternate Results for Uniform Temperature Case. To gain some insight into the sensitivity of the results with respect to the assumed profile shapes, calculations also were performed for the following velocity and temperature profiles

$$\frac{U}{U_1} = \frac{Y}{\Delta}\left(1 - \frac{Y}{\Delta}\right)^4 \ldots\ldots [21]$$

$$\frac{\theta}{\theta_w} = \left(1 - \frac{Y}{\Delta}\right)^2 \ldots\ldots [22]$$

The results are summarized as follows:

The solution dependent on time only is

$$\delta = \sqrt{(12)}\,\tau^{1/2} \ldots\ldots [23]$$

$$u_1 = \frac{4\mathrm{Ra}_L\mathrm{Pr}}{0.6 + \mathrm{Pr}}\tau \quad [24]$$

The steady-state solution is

$$\delta = 4.60\,(0.377 + \mathrm{Pr})^{1/4}(\mathrm{Ra}_L\mathrm{Pr})^{-1/4}x^{1/4} \quad [25]$$

$$u_1 = 7.06\,(0.377 + \mathrm{Pr})^{-1/2}(\mathrm{Ra}_L\mathrm{Pr})^{1/2}x^{1/2} \quad [26]$$

The equation of the I$_+$° line, which gives the time at which one-dimensional diffusion ends at each plate position, is

$$\tau = 2.48\,(0.6 + \mathrm{Pr})^{1/2}(\mathrm{Ra}_L\mathrm{Pr})^{-1/2}x^{1/2} \quad [27]$$

The equation of the I$_-$° line, which gives the time at which steady state is reached at each position along the plate, is

$$\tau = 7.10\,(0.377 + \mathrm{Pr})^{1/2}(\mathrm{Ra}_L\mathrm{Pr})^{-1/2}x^{1/2} \quad [28]$$

Comparisons With Exact Solutions. Some insight into the accuracy of the integral method can be obtained by comparing the velocity and temperature profiles obtained in the particular solutions with more precise solutions of the differential equations for free convection. As mentioned previously, in the initial stages of motion, heat is conducted into the fluid in a one-dimensional fashion, and hence the temperature profile during this period is the same as that resulting from suddenly raising the surface temperature of a semi-infinite solid. Then from Jakob[7] we have

$$\Phi = \frac{\theta}{\theta_w} = 1 - \mathrm{erf}\,\eta, \text{ where } \eta = \frac{Y}{2\sqrt{(\alpha T)}} \quad [29]$$

From Illingworth,[8] the velocity profile for this stage of the motion is for Pr = 1

$$U = 2g T\beta\theta_w\eta\left[\frac{1}{\sqrt{\pi}}e^{-\eta^2} + \eta\,\mathrm{erf}\,\eta - \eta\right] \quad [30]$$

In the integral method the approximate temperature profile in the initial stage of the motion is obtained by substituting Δ from Equation [13] or [23] into the profile Equation [4] or [22] with the result

[7] Reference (7), p. 253.
[8] Reference (1), p. 612.

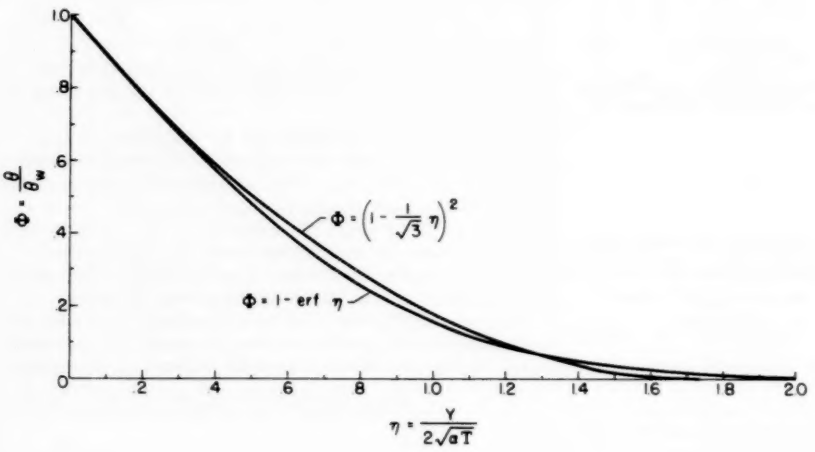


Fig. 2 Dimensionless Temperature Profiles During Initial One-Dimensional Transient-Flow Development on a Plate at Uniform Temperature

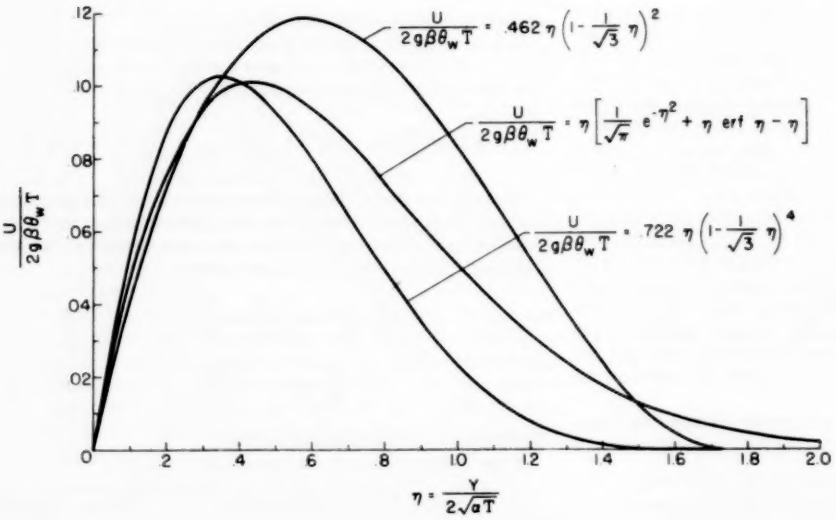


Fig. 3 Dimensionless Velocity Profiles During Initial One-Dimensional Transient-Flow Development on a Plate at Uniform Temperature. Pr = 1

$$\Phi = \left(1 - \frac{1}{\sqrt{3}}\,\eta\right)^3$$

This is compared with the exact solution, Equation [29], in Fig. 2.

The velocity profiles for the integral method can be derived by substituting U_1 from Equation [14] and Δ from Equation [13] into Equation [3] with the result

$$\frac{U}{2g\beta\theta_w T} = \frac{2}{(1.5 + \mathrm{Pr})}\left(\frac{1}{\sqrt{3}}\,\eta\right)\left(1 - \frac{1}{\sqrt{3}}\,\eta\right)^2 \quad \ldots [31]$$

or by substituting Equations [24] and [23] into Equation [21] with the result

$$\frac{U}{2g\beta\theta_w T} = \frac{2}{(0.6 + \mathrm{Pr})}\left(\frac{1}{\sqrt{3}}\,\eta\right)\left(1 - \frac{1}{\sqrt{3}}\,\eta\right)^4 \quad \ldots [32]$$

These are evaluated for Pr = 1 and compared with Equation [30] in Fig. 3.

In a similar fashion the steady-state velocity and temperature profiles derived in the integral method can be compared with the exact solution given in (8). By substituting the steady-state solution, Equations [17] and [18], into Equations [3] and [4], the following profiles are obtained

$$\frac{\frac{UX}{\nu}}{2\,(\mathrm{Gr}_X)^{1/2}} = \frac{0.930\ \mathrm{Pr}^{1/2}}{(0.952 + \mathrm{Pr})^{3/4}}\,\chi\left[1 - \frac{\sqrt{(2)}\ \mathrm{Pr}^{1/2}}{3.93\,(0.952 + \mathrm{Pr})^{1/4}}\,\chi\right]^3 \quad \ldots [33]$$

$$\Phi = \left[1 - \frac{\sqrt{(2)}\ \mathrm{Pr}^{1/2}}{3.93\,(0.952 + \mathrm{Pr})^{1/4}}\,\chi\right]^2 \quad \ldots [34]$$

where

$$\chi = \left(\frac{\mathrm{Gr}_X}{4}\right)^{1/4}\frac{Y}{X}$$

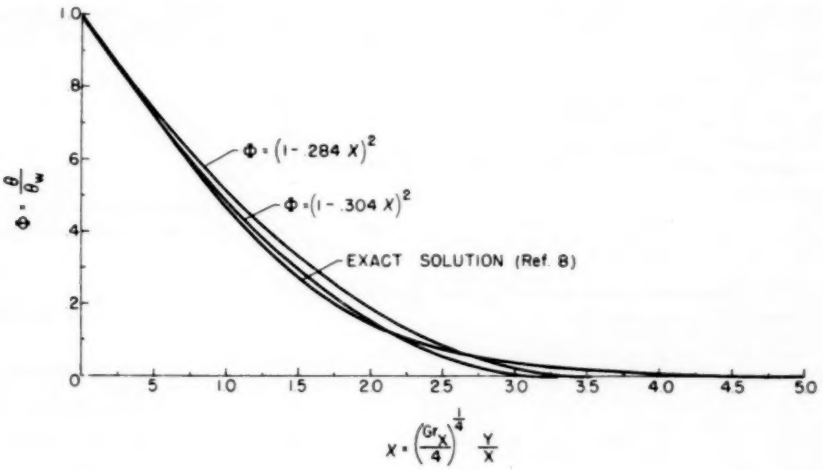


FIG. 4 DIMENSIONLESS TEMPERATURE PROFILES FOR STEADY-STATE FREE CONVECTION ON A CONSTANT-TEMPERATURE PLATE. PR = 1

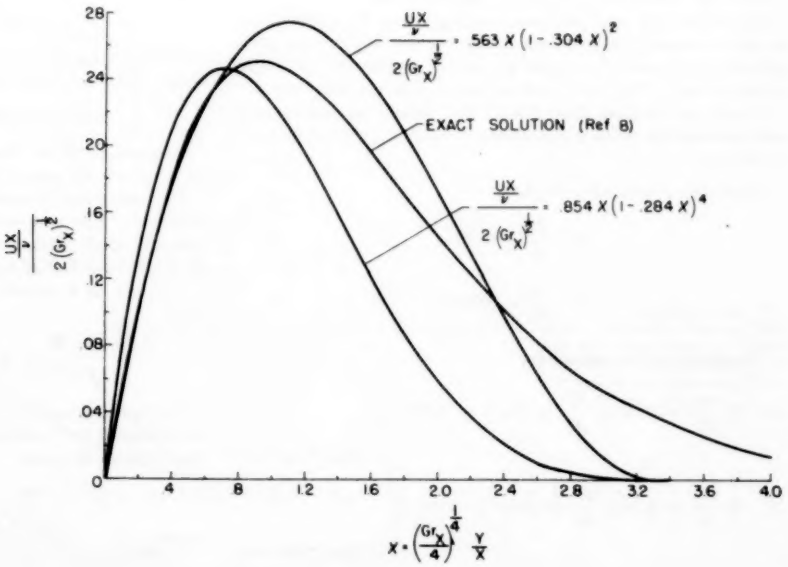


FIG. 5 DIMENSIONLESS VELOCITY PROFILES FOR STEADY-STATE FREE CONVECTION ON A CONSTANT-TEMPERATURE PLATE. PR = 1

If Equations [25] and [26] are substituted into Equations [21] and [22] the resulting profiles are

$$\frac{UX}{\nu} = \frac{1.085\ \mathrm{Pr}^{1/2}}{(0.377 + \mathrm{Pr})^{1/4}}\chi\left[1 - \frac{\sqrt{(2)}\ \mathrm{Pr}^{1/2}}{4.60\ (0.377 + \mathrm{Pr})^{1/4}}\chi\right]^2 \quad [35]$$

$$\Phi = \left[1 - \frac{\sqrt{(2)}\ \mathrm{Pr}^{1/2}}{4.60\ (0.377 + \mathrm{Pr})^{1/4}}\chi\right]^2 \quad [36]$$

These approximations are compared with the exact solutions in Figs. 4 and 5.

Plate at Uniform Heat Flux

Basic Equations. This problem is essentially the same as the previous case for uniform wall temperature, except that now the plate, initially at ambient temperature, suddenly begins to produce a uniform heat flux at its surface. In this instance the surface temperature will vary with both time and position on the plate, and hence the equations of motion and energy are first rearranged so that q, the uniform heat output per unit area, will replace the now variable θ_w.

As a first example it is assumed that the profiles of Equations [3] and [4] also can be applied for the uniform heat-flux case. The profiles are inserted into Equations [1] and [2], and the integration carried out with θ_w variable. θ_w is eliminated by noting that

$$q = -k\left.\frac{\partial\theta}{\partial Y}\right|_{Y=0} = 2\frac{k}{\Delta}\theta_w, \quad \text{or} \quad \theta_w = \frac{q\Delta}{2k} \quad [37]$$

and the equations of motion and energy then become, in dimensionless form

$$\frac{1}{12\mathrm{Pr}}\frac{\partial}{\partial\tau}(u_1\delta) + \frac{1}{105\mathrm{Pr}}\frac{\partial}{\partial x}(\delta u_1^2) = \frac{1}{6}\mathrm{Ra}_L^*\delta^2 - \frac{u_1}{\delta} \quad [38]$$

$$\frac{\partial}{\partial\tau}(\delta^2) + \frac{1}{10}\frac{\partial}{\partial x}(u_1\delta^2) = 6 \quad [39]$$

where Ra_L^* is a modified Rayleigh number based on q.

Equations [38] and [39] are rearranged into the form of the equations treated in Appendix 1, and the solution for the characteristics follows in the same manner as for the constant-temperature case. The results are summarized in the following.

Results for Uniform Heat Flux. The results for the velocity and temperature profiles, Equations [3] and [4], are summarized as follows:

Equations of characteristic lines

$$\mathrm{I}_+ \quad \frac{dx}{d\tau} = 0.187u_1$$

$$\mathrm{I}_- \quad \frac{dx}{d\tau} = 0.0917u_1$$

The characteristic equations

$$\mathrm{II}_+ \quad \delta\frac{du_1}{d\tau} + 0.167u_1\frac{d\delta}{d\tau} + (2.50 + 12\mathrm{Pr})\frac{u_1}{\delta} - 2\mathrm{Ra}_L^*\mathrm{Pr}\delta^2 = 0$$

$$\mathrm{II}_- \quad \delta\frac{du_1}{d\tau} - 1.74u_1\frac{d\delta}{d\tau} + (8.21 + 12\mathrm{Pr})\frac{u_1}{\delta} - 2\mathrm{Ra}_L^*\mathrm{Pr}\delta^2 = 0$$

The solution dependent on time only is

$$\delta = \sqrt{(6)}\,\tau^{1/2} \quad [40]$$

$$u_1 = \frac{\sqrt{(6)}\ \mathrm{Ra}_L^*\mathrm{Pr}}{1 + \mathrm{Pr}}\tau^{3/2} \quad [41]$$

The steady-state solution is

$$\delta = 3.25\ (0.800 + \mathrm{Pr})^{1/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-1/4}x^{1/4} \quad [42]$$

$$u_1 = 5.70\ (0.800 + \mathrm{Pr})^{-2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{2/4}x^{2/4} \quad [43]$$

The equation of the $\mathrm{I}_+{}^\circ$ line is

$$\tau = 1.97\ (1 + \mathrm{Pr})^{2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-2/4}x^{2/4} \quad [44]$$

The time at which steady state is reached is given by the $\mathrm{I}_-{}^\circ$ line

$$\tau = 4.78\ (0.8 + \mathrm{Pr})^{2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-2/4}x^{2/4} \quad [45]$$

The uniform heat-flux case also was evaluated for another set of profiles to obtain an indication of the sensitivity of the results to the profiles chosen. Using the profiles

$$\frac{U}{U_1} = \frac{Y}{\Delta}\left(1 - \frac{Y}{\Delta}\right)^2 \quad [46]$$

$$\frac{\theta}{\theta_w} = \left(1 - \frac{Y}{\Delta}\right)^3 \quad [47]$$

the following results were obtained:

The solution dependent on time only is

$$\delta = \sqrt{(12)}\,\tau^{1/3} \quad [48]$$

$$u_1 = \sqrt{(12)}\,\frac{\mathrm{Ra}_L^*\mathrm{Pr}}{2 + \mathrm{Pr}}\tau^{2/3} \quad [49]$$

The steady-state solution is

$$\delta = 4.33\ (1.68 + \mathrm{Pr})^{1/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-1/4}x^{1/4} \quad [50]$$

$$u_1 = 6.73\ (1.68 + \mathrm{Pr})^{-2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{2/4}x^{2/4} \quad [51]$$

The equation of the $\mathrm{I}_+{}^\circ$ line is

$$\tau = 1.71\ (2 + \mathrm{Pr})^{2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-2/4}x^{2/4} \quad [52]$$

The equation of the $I_-{}^\circ$ line which gives the time required for steady state to be reached is

$$\tau = 4.33\ (1.68 + \mathrm{Pr})^{2/4}(\mathrm{Ra}_L^*\mathrm{Pr})^{-2/4}x^{2/4} \quad [53]$$

Comparison With Exact Solutions. A comparison can now be made of the velocity and temperature profiles for the particular solutions obtained by the integral method with exact solutions. The temperature profile applicable during the initial pure diffusion portion of the motion is obtained from Jakob[9] as being the result of applying a source of uniform heat flux to the surface of a semi-infinite solid. This can be put into the form

$$\frac{\theta k}{q\sqrt{(\alpha T)}} = 2\eta\left[\mathrm{erf}\ \eta + \frac{e^{-\eta^2}}{\sqrt{(\pi)}\eta} - 1\right] \quad [54]$$

The integral method profiles for the two examples calculated are obtained by inserting Equation [40] into the profile Equation [4] with the result

$$\frac{\theta k}{q\sqrt{(\alpha T)}} = \frac{\sqrt{6}}{2}\left(1 - \frac{2}{\sqrt{6}}\eta\right)^2 \quad [55]$$

[9] Reference (7), p. 258.

or by substituting Equation [48] into Equation [47] with the result

$$\frac{\theta k}{q\sqrt{(\alpha T)}} = \frac{\sqrt{12}}{3}\left(1 - \frac{1}{\sqrt{3}}\,\eta\right)^3 \quad \ldots\ldots\ldots [56]$$

These three profiles are shown in Fig. 6.

The velocity profile in the early stage of the motion for the uniform heat-flux case and Pr = 1 is derived in Appendix 2

$$\frac{Uk}{g\beta q\sqrt{(\alpha T)}T} = 2\eta\left[\frac{1}{2} + \eta^2 - \frac{\eta}{\sqrt{\pi}}\,e^{-\eta^2} - \frac{1}{2}\,\text{erf}\,\eta - \eta^2\,\text{erf}\,\eta\right] \quad \ldots\ldots\ldots [57]$$

The corresponding integral-method profiles are obtained by substituting Equations [40] and [41] into Equation [3] with the result

$$\frac{Uk}{g\beta q\sqrt{(\alpha T)}T} = \frac{2}{(1 + \text{Pr})}\,\eta\left(1 - \frac{2}{\sqrt{6}}\,\eta\right)^2 \quad \ldots\ldots [58]$$

or by inserting Equations [48] and [49] into Equation [46] with the result

$$\frac{Uk}{g\beta q\sqrt{(\alpha T)}T} = \frac{2}{(2 + \text{Pr})}\,\eta\left(1 - \frac{1}{\sqrt{3}}\,\eta\right)^2 \quad \ldots\ldots [59]$$

These three velocity profiles are compared in Fig. 7.

The steady-state velocity and temperature profiles derived by the integral method can be compared with the exact solutions in (9). The approximate profiles are obtained by substituting the steady-state solution, Equations [42] and [43] into Equations [3] and [4] with the results

$$\frac{\dfrac{UX}{\nu}}{5\left(\dfrac{\text{Gr}_X^*}{5}\right)^{2/5}} = \frac{0.922\ \text{Pr}^{1/5}\chi^*}{(0.800 + \text{Pr})^{2/5}}\left[1 - \frac{0.425\ \text{Pr}^{2/5}}{(0.800 + \text{Pr})^{1/5}}\,\chi^*\right]^2 \quad \ldots\ldots [60]$$

$$\Phi = \left[1 - \frac{0.425\ \text{Pr}^{2/5}}{(0.800 + \text{Pr})^{1/5}}\,\chi^*\right]^2 \quad \ldots\ldots\ldots [61]$$

or by substituting Equations [50] and [51] into Equations [46] and [47] with the results

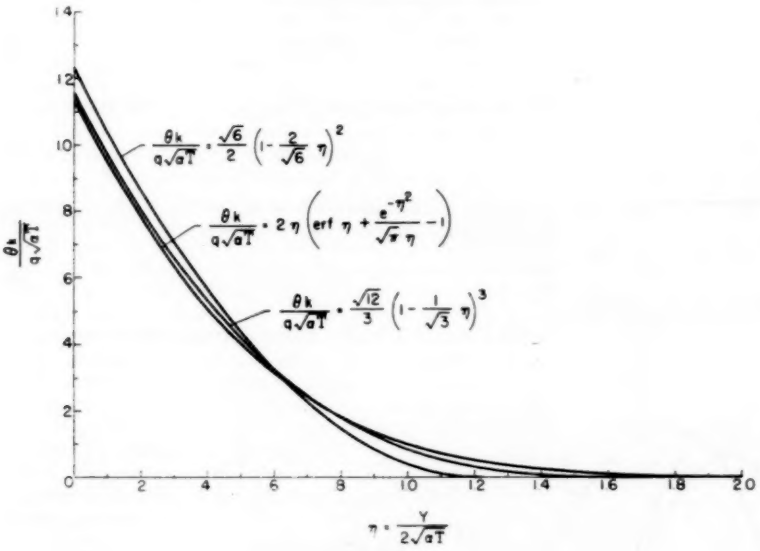


FIG. 6 DIMENSIONLESS TEMPERATURE PROFILES DURING INITIAL ONE-DIMENSIONAL TRANSIENT-FLOW DEVELOPMENT FOR A PLATE WITH UNIFORM HEAT FLUX

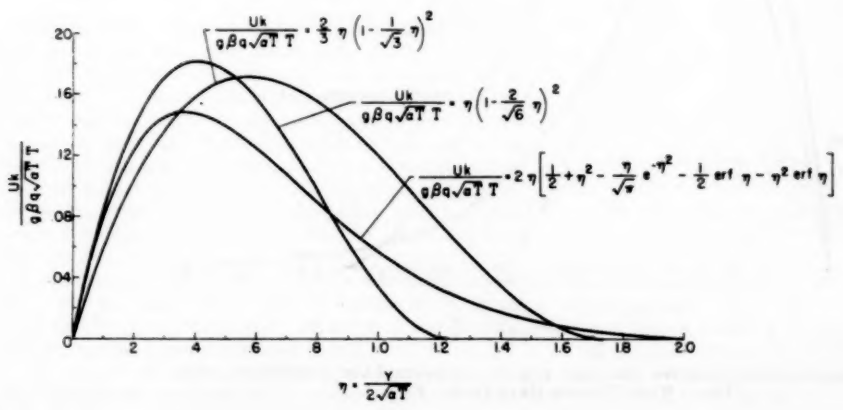


FIG. 7 DIMENSIONLESS VELOCITY PROFILES DURING INITIAL ONE-DIMENSIONAL TRANSIENT-FLOW DEVELOPMENT ON A PLATE WITH UNIFORM HEAT FLUX. PR = 1

$$\frac{\frac{UX}{\nu}}{5\left(\frac{Gr_x^*}{5}\right)^{3/4}} = \frac{0.816\ Pr^{1/4}}{(1.68 + Pr)^{3/4}}\chi^* \left[1 - \frac{0.319\ Pr^{3/4}\chi^*}{(1.68 + Pr)^{1/4}}\right]^2 \quad [62]$$

$$\Phi_* = \left[1 - \frac{0.319\ Pr^{3/4}}{(1.68 + Pr)^{1/4}}\chi^*\right]^3 \quad [63]$$

where

$$\chi^* = \frac{Y}{X}\left(\frac{Gr_X^*}{5}\right)^{1/4}$$

These profiles are plotted for Pr = 1 in Figs. 8 and 9.

Heat-Transfer Coefficients

For steady-state conditions exact and approximate heat-transfer coefficients are available in (5, 8, 9, and 10) for both uniform temperature and uniform heat-flux boundary conditions and these results will not be presented here.

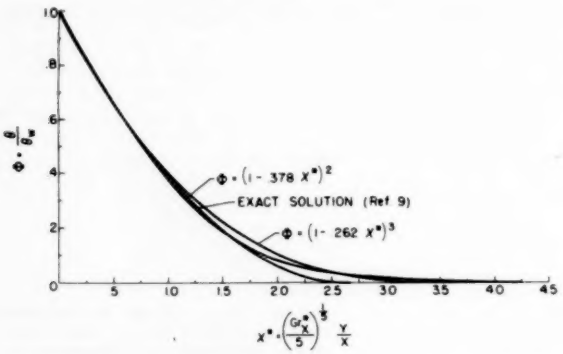


Fig. 8 Dimensionless Temperature Profiles for Steady-State Free Convection From a Plate With Uniform Heat Flux. Pr = 1

For the initial one-dimensional transient, the heat-transfer coefficient can be found from the relation

$$h = -\frac{1}{\theta_w}\, k \left.\frac{\partial \theta}{\partial Y}\right|_{Y=0}$$

For the constant-wall-temperature case the two examples calculated yield the same result which can be obtained by differentiating the profile Equation [4] and substituting δ from Equation [13]. This gives

$$h = 0.578\ \frac{k}{(\alpha T)^{1/2}} \quad [64]$$

The comparable exact solution can be found by evaluating the temperature gradient at the wall from Equation [29]. This results in

$$h = 0.564\ \frac{k}{(\alpha T)^{1/2}} \quad [65]$$

In a similar fashion heat-transfer coefficients during the initial one-dimensional transient can be evaluated for the case of uniform wall heat flux. For the first example calculated using a parabolic temperature distribution, the result is

$$h = 0.816\ \frac{k}{(\alpha T)^{1/2}} \quad [66]$$

For the cubic temperature profile the coefficient is given by

$$h = 0.866\ \frac{k}{(\alpha T)^{1/2}} \quad [67]$$

If the temperature derivative at the wall is evaluated from the exact solution, Equation [54], the result is

$$h = 0.886\ \frac{k}{(\alpha T)^{1/2}} \quad [68]$$

which compares favorably with the integral-method results.

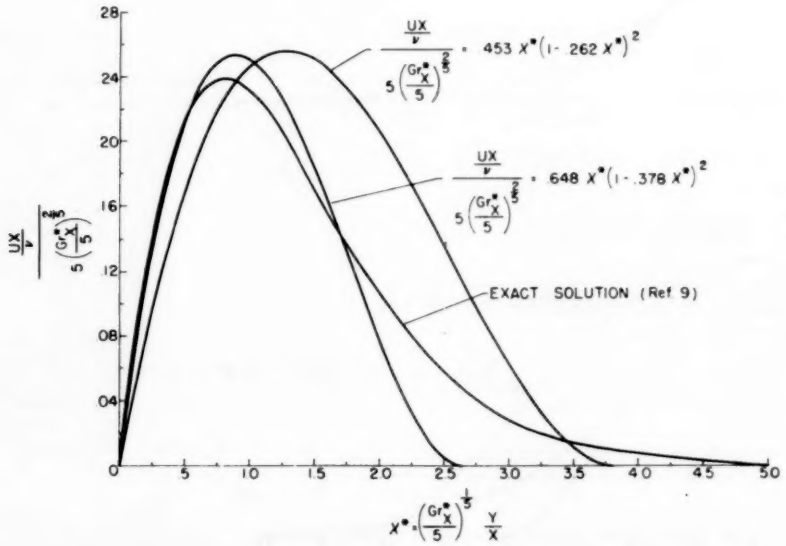


Fig. 9 Dimensionless Velocity Profiles for Steady-State Free Convection From a Plate With Uniform Heat Flux. Pr = 1

Discussion

The velocity and temperature profiles for a plate suddenly raised to a uniform temperature are compared with exact solutions in Figs. 2, 3, 4, and 5. The exact profiles for the initial one-dimensional transient portion of the flow are quite similar in shape to the steady-state curves, so the assumption in the integral method that profiles remain similar in shape throughout the transient appears reasonable. The approximate temperature profiles are in quite good agreement with the exact solutions, while the velocity profiles show a larger deviation. However, the approximate velocity curves fall on either side of the exact profiles and hence a better fit to the velocity curves would be expected to yield results between the values computed. Since the two constant-temperature computations yield times required to reach steady state, which are in agreement within from about 16 to 30 per cent for all Prandtl numbers, the computations are evidently not highly sensitive to the velocity profile. The recommended equation given in the next section is taken as an average of the two computations presented.

The profiles for the uniform heat-flux case are given in Figs. 6, 7, 8, and 9. The temperature profiles are again in good agreement with the exact solutions, while the velocity profiles resulting from the calculation using Equations [3] and [4] are better approximations than those resulting from Equations [46] and [47]. The results of the former calculation are therefore recommended with the results of the latter being used to gain an indication of the sensitivity of the results to the profile shapes used. A small sensitivity is indicated by comparing Equations [45] and [53] which yield times required to reach steady state which agree within 20 per cent for all Prandtl numbers.

The results indicate that an increase in $Ra_L Pr$ or $Ra_L{}^* Pr$ will cause a decrease in the value of τ required to reach steady state. Thus less time is required to develop the convective flow if any of the factors, g, β, q, or θ_w are increased. This agrees with the results of (2), where it is stated for the constant-wall-temperature case, that the convection process proceeds more quickly and violently as the wall-to-fluid temperature difference is increased.

The analysis indicates that during the transient-flow development the boundary-layer thickness exceeds for a time the steady-state value. This may be illustrated by use of the first computation presented for the constant-temperature boundary condition. If Equations [13] and [19] are combined, the boundary-layer thickness is found as a function of x at the termination of the one-dimensional diffusion portion of the transient

$$\delta = 4.65\,(1.5 + \mathrm{Pr})^{1/4}\,(\mathrm{Ra}_L \mathrm{Pr})^{-1/4}\,x^{1/4}$$

For Pr = 1, this yields a boundary-layer thickness about 25 per cent larger than the steady-state value from Equation [17]. This indicates that, at a given position along the plate, the heat-conduction process is evidently sufficiently rapid to enable the boundary layer to grow beyond the steady-state thickness before the constraints introduced by the leading edge can propagate to that location and prevent the growth from continuing as if the plate were infinite in length. The transient boundary-layer growth at two successive times is illustrated by Fig. 10 which has been calculated from Equations [13], [17], [19], and [20]. The intermediate region of adjustment has been faired in as a dotted line. For each time, this shows the portions of the boundary layer which are at steady state, adjusting to steady state, or still undergoing a one-dimensional growth.

The overgrowth in the boundary layer causes a minimum in the heat-transfer coefficient as illustrated in Fig. 11 where the heat-transfer coefficient is plotted as a function of time for two positions along the plate. The curves have been calculated from Equations [64], [19], and [20] with the steady-state values taken from Eckert.[10] It is not known if this minimum is a consequence of the approximate method utilized in the analysis or if

[10] Reference (5), p. 162.

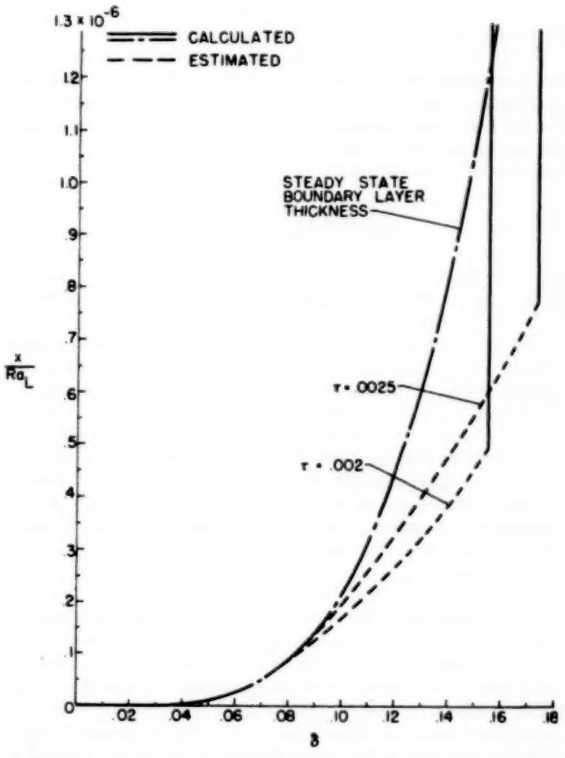


Fig. 10 Transient Boundary-Layer Growth for a Plate at Uniform Temperature

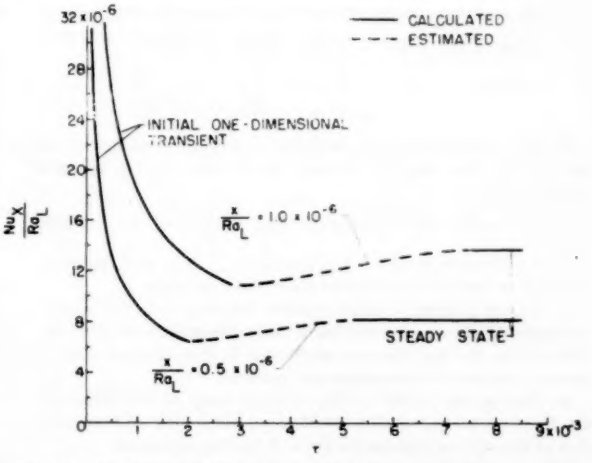


Fig. 11 Transient Variation of Heat-Transfer Coefficient on a Plate at Uniform Temperature. Pr = 1

it is a physical reality, and experimental information is needed for comparison.

To the author's knowledge the only other work on transient free convection from a plate of semi-infinite length is presented in (2) where a method of successive approximations is employed. For the first approximation, convective effects are neglected, and the energy equation is reduced to the two-dimensional transient heat-conduction equation

$$\frac{\partial t}{\partial T} = \alpha\left(\frac{\partial^2 t}{\partial X^2} + \frac{\partial^2 t}{\partial Y^2}\right)$$

This was solved by an approximate method, and for distances sufficiently far from the leading edge the results are independent of X as expected. However, in this one-dimensional region the profile is somewhat steeper than the one-dimensional solution, Equation [29], and hence there appears to be an error in the numerical work. The first approximation is used in the original differential equations to obtain an improved solution. The results of the second approximation are only presented for very short times so it is not possible to evaluate how large a change the second approximation could introduce as time increases. Since the results are thus confined to the initial stages of the motion, it is not possible to make a comparison with the present work as to the time required to achieve steady state.

Summary of Results

1 The solution of the transient free-convection equations of momentum and energy, placed in integral form (Karman-Pohlhausen method), yields two families of characteristic lines on the τ-x plane. The two characteristics passing through the origin divide the plane into three regions: (a) A region of initial one-dimensional boundary-layer growth, (b) a region of readjustment in which the flow is influenced by the leading edge, and (c) a steady-state region which is reached at a different time for each location along the plate.

2 For a plate suddenly raised to uniform temperature, the time required to reach steady state is given approximately by

$$\tau = \left[\frac{5.24\,(0.952 + \mathrm{Pr})^{1/2} + 7.10\,(0.377 + \mathrm{Pr})^{1/2}}{2}\right](\mathrm{Ra}_L\mathrm{Pr})^{-1/2}x^{1/2} \quad \ldots\ldots [69]$$

3 For a plate suddenly producing a uniform heat flux at its surface, the time required to reach steady state is given approximately by

$$\tau = 4.78\,(0.8 + \mathrm{Pr})^{2/5}(\mathrm{Ra}_L{}^*\mathrm{Pr})^{-2/5}x^{2/5} \quad \ldots\ldots [70]$$

4 An increase in any of the quantities g, β, q, or θ_w causes a decrease in the time required to achieve steady state.

5 In the process of transient-flow development, the boundary-layer thickness exceeds for a time the steady-state value. This causes the heat-transfer coefficient to pass through a minimum before steady conditions are reached.

6 During the initial one-dimensional stage of the transient, the free-convection velocity profile on a plate with uniform heat flux at the surface is given for Pr = 1, by the expression

$$\frac{Uk}{g\beta q\sqrt{(\alpha T)T}} = 2\eta\left[\frac{1}{2} + \eta^2 - \frac{\eta}{\sqrt{\pi}}e^{-\eta^2} - \frac{1}{2}\,\mathrm{erf}\,\eta - \eta^2\,\mathrm{erf}\,\eta\right]$$

BIBLIOGRAPHY

1 "Unsteady Laminar Flow of Gas Near an Infinite Flat Plate," by C. R. Illingworth, Proceedings of the Cambridge Philosophical Society, vol 46, part 4, October, 1950, pp. 603–613.

2 "The Heat Transfer by Natural Convection in the Unsteady State on a Vertical Flat Wall," by S. Sugawara and I. Michiyoshi, Proceedings of the First Japan National Congress for Applied Mechanics, 1951, National Committee for Theoretical and Applied Mechanics, Science Council of Japan, May, 1952, pp. 501–506.

3 "Modern Developments in Fluid Dynamics," edited by S. Goldstein, Oxford University Press, London, England, 1938.

4 "Calculation of Unsteady Boundary Layers in Two-Dimensional Laminar Flow," by H. Schuh, KTH Aeronautical Rapport FL 141, Flygtekniska Laboratoriet, Stockholm, Sweden, 1953.

5 "Introduction to the Transfer of Heat and Mass," by E. R. G. Eckert, McGraw-Hill Book Company, Inc., New York, N. Y., 1950.

6 "Supersonic Flow and Shock Waves," by R. Courant and K. O. Friedrichs, Interscience Publishers, Inc., New York, N. Y., 1948.

7 "Heat Transfer," by M. Jakob, John Wiley & Sons, Inc., New York, N. Y., 1949.

8 "An Analysis of Laminar Free-Convection Flow and Heat Transfer About a Flat Plate Parallel to the Direction of the Generating Body Force," by S. Ostrach, NACA TR 1111, 1953.

9 "Laminar Free Convection From a Vertical Plate With Uniform Surface Heat Flux," by E. M. Sparrow and J. L. Gregg, Trans. ASME, vol. 78, 1956, pp. 435–440.

10 "Analysis of Laminar and Turbulent Free Convection From a Smooth Vertical Plate With Uniform Heat Dissipation Per Unit Surface Area," by R. Siegel, General Electric Report, R54GL89, April, 1954.

Appendix 1

Derivation of Characteristic Equations

If the notation of (6), section 21, is adopted, Equations [5] and [6] can be put into the general form

$$A_1\frac{\partial u_1}{\partial \tau} + B_1\frac{\partial u_1}{\partial x} + C_1\frac{\partial \delta}{\partial \tau} + D_1\frac{\partial \delta}{\partial x} + E_1 = 0 \ldots [71]$$

$$A_2\frac{\partial u_1}{\partial \tau} + B_2\frac{\partial u_1}{\partial x} + C_2\frac{\partial \delta}{\partial \tau} + D_2\frac{\partial \delta}{\partial x} + E_2 = 0 \ldots [72]$$

where the coefficients are

$$A_1 = \frac{\delta}{12\mathrm{Pr}};\quad B_1 = \frac{2\delta u_1}{105\mathrm{Pr}};\quad C_1 = \frac{u_1}{12\mathrm{Pr}};\quad D_1 = \frac{u_1{}^2}{105\mathrm{Pr}};\quad E_1 = -\frac{1}{3}\mathrm{Ra}_L\delta + \frac{u_1}{\delta}$$

$$A_2 = 0;\quad B_2 = \frac{\delta}{10};\quad C_2 = 1;\quad D_2 = \frac{u_1}{10};\quad E_2 = -\frac{6}{\delta}$$

To determine whether Equations [71] and [72] are hyperbolic, parabolic, or elliptic, the sign of the quantity $ac - b^2$ is examined, where

$$a = A_1C_2 - A_2C_1 = \frac{\delta}{12\mathrm{Pr}}$$

$$b = \frac{1}{2}(A_1D_2 - A_2D_1 + B_1C_2 - B_2C_1) = \frac{2\delta u_1}{105\mathrm{Pr}}$$

$$c = B_1D_2 - B_2D_1 = \frac{\delta u_1{}^2}{1050\mathrm{Pr}}$$

Then

$$ac - b^2 = \frac{1}{105}\,\frac{\delta^2 u_1}{\mathrm{Pr}^2}\left[\frac{1}{120} - \frac{1}{105}\right]$$

is less than zero which indicates that the set of equations is hyperbolic and a solution can be obtained by the method of characteristics. As shown in (6) the equation for the characteristic lines is given by

$$a\left(\frac{dx}{d\tau}\right)^2 - 2b\frac{dx}{d\tau} + c = 0$$

Substituting for a, b, and c, and solving for $(dx)/(d\tau)$, we obtain

$$\frac{dx}{d\tau} = \frac{4}{35} u_1 \left(1 \pm \frac{\sqrt{2}}{4}\right)$$

The two families of characteristic lines will be designated for convenience by I_+ and I_-

$$I_+ \quad \frac{dx}{d\tau} = \frac{4}{35} u_1 \left(1 + \frac{\sqrt{2}}{4}\right) = 0.155\, u_1 \ldots\ldots [73]$$

$$I_- \quad \frac{dx}{d\tau} = \frac{4}{35} u_1 \left(1 - \frac{\sqrt{2}}{4}\right) = 0.0739\, u_1 \ldots\ldots [74]$$

This shows that the slopes of the characteristic lines are proportional to the velocity within the free-convection boundary layer.

The characteristic equations which are valid along the characteristic lines can be obtained from the relation

$$T\left(\frac{du_1}{d\tau}\right) + \left(a\frac{dx}{d\tau} - S\right)\frac{d\delta}{d\tau} + \left(K\frac{dx}{d\tau} - H\right) = 0 \ldots [75]$$

where

$$T = A_1B_2 - A_2B_1 = \frac{\delta^2}{120\mathrm{Pr}}$$

$$S = B_1C_2 - B_2C_1 = \frac{3}{280}\frac{\delta u_1}{\mathrm{Pr}}$$

$$K = A_1E_2 - A_2E_1 = -\frac{1}{2\mathrm{Pr}}$$

$$H = B_1E_2 - B_2E_1 = \left(\frac{1}{3}\mathrm{Ra}_L\delta - \frac{u_1}{\delta}\right)\frac{\delta}{10} - \frac{4}{35}\frac{u_1}{\mathrm{Pr}}$$

These equations are inserted into Equation [75] and $(dx)/(d\tau)$ is eliminated using Equation [73] and then Equation [74] to yield the two characteristic equations

$$II_+ \quad \delta\frac{du_1}{d\tau} + 0.261u_1\frac{d\delta}{d\tau} + (4.43 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\mathrm{Pr}\delta = 0 \ldots\ldots [76]$$

$$II_- \quad \delta\frac{du_1}{d\tau} - 0.547u_1\frac{d\delta}{d\tau} + (9.28 + 12\mathrm{Pr})\frac{u_1}{\delta} - 4\mathrm{Ra}_L\mathrm{Pr}\delta = 0 \ldots\ldots [77]$$

Appendix 2

Velocity Profile for Initial One-Dimensional Stage of Motion With Uniform Heat Flux

The boundary-layer equations for transient free convection can be obtained by adding the time-dependent terms to the steady-state equations as given for example by Jakob.[11] The equations of momentum, continuity, and energy are then

$$\frac{\partial U}{\partial T} + U\frac{\partial U}{\partial X} + V\frac{\partial U}{\partial Y} = \nu\frac{\partial^2 U}{\partial Y^2} + g\beta\theta \ldots\ldots [78]$$

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} = 0 \ldots\ldots [79]$$

$$\frac{\partial \theta}{\partial T} + U\frac{\partial \theta}{\partial X} + V\frac{\partial \theta}{\partial Y} = \alpha\frac{\partial^2\theta}{\partial Y^2} \ldots\ldots [80]$$

The portion of the plate considered is that which has not yet been influenced by the leading edge and thus has a flow and temperature distribution independent of X. Then

$$\frac{\partial U}{\partial X} = \frac{\partial \theta}{\partial X} = 0$$

From Equation [79], $\partial V/\partial Y = 0$, and since $V = 0$ at $Y = 0$, it follows that $V = 0$ for all Y. Equations [78] and [80] then reduce to

$$\frac{\partial U}{\partial T} = \nu\frac{\partial^2 U}{\partial Y^2} + g\beta\theta \ldots\ldots [81]$$

$$\frac{\partial \theta}{\partial T} = \alpha\frac{\partial^2\theta}{\partial Y^2} \ldots\ldots [82]$$

The solution to Equation [82], subject to the boundary condition of suddenly applying uniform heat flux, is given as[12]

$$\theta = \frac{2q}{k}\sqrt{(\alpha T)}\eta\left(\mathrm{erf}\,\eta + \frac{e^{-\eta^2}}{\sqrt{(\pi)}\eta} - 1\right), \text{ where } \eta = \frac{Y}{2\sqrt{(\alpha T)}}$$

This is substituted into Equation [81] and the result is transformed into an ordinary differential equation by letting

$$U = \frac{2g\beta q\sqrt{\alpha}}{k} T^{3/2}\zeta(\eta) \ldots\ldots [83]$$

This yields

$$\frac{d^3\zeta}{d\eta^3} + \frac{2}{\mathrm{Pr}}\eta\frac{d\zeta}{d\eta} - \frac{6}{\mathrm{Pr}}\zeta = \frac{4}{\mathrm{Pr}}\eta\left(-\mathrm{erf}\,\eta - \frac{e^{-\eta^2}}{\sqrt{(\pi)}\eta} + 1\right). [84]$$

The general solution to Equation [84] for the case of $\mathrm{Pr} = 1$ was found in part from a power-series solution and in part by substituting trial functions into the differential equation until a particular integral was found. The result was

$$\zeta = A\left(\eta + \frac{2}{3}\eta^3\right) + B\left(e^{-\eta^2} + \eta^2 e^{-\eta^2} + \frac{3}{2}\sqrt{(\pi)}\eta\,\mathrm{erf}\,\eta + \sqrt{(\pi)}\eta^3\,\mathrm{erf}\,\eta\right) + \eta\,\mathrm{erf}\,\eta - \eta + \frac{1}{\sqrt{\pi}}e^{-\eta^2}$$

The arbitrary constants A and B are evaluated from the boundary conditions that $\zeta = 0$ at $\eta = 0$ and $\eta = \infty$. This results in

$$A = \frac{3}{2}; \quad B = -\frac{1}{\sqrt{\pi}}$$

With these constants the solution can be simplified to the final result valid for $\mathrm{Pr} = 1$

$$U = \frac{2g\beta q\sqrt{(\alpha T)}}{k} T\eta\left[\frac{1}{2} + \eta^2 - \frac{\eta}{\sqrt{\pi}}e^{-\eta^2} - \frac{1}{2}\mathrm{erf}\,\eta - \eta^2\,\mathrm{erf}\,\eta\right] \ldots\ldots [85]$$

It is noted that the result also satisfies the boundary condition that $U = 0$ at $T = 0$ for all Y.

[11] Reference (7), p. 446.

[12] Reference (7), p. 258.

Discussion

R. J. GOLDSTEIN.[13] The question of calculating transient thermal boundary-layer growth is an interesting one which unfortunately has received little attention. The author is to be congratulated on his novel application of the method of characteristics to this problem.

There was some doubt in the writer's mind as to the physical phenomena or characteristic of the differential equations causing the dip in heat-transfer coefficient following the step-function rise in temperature. An attempt was made to investigate a possible source of error by using better approximations of the temperature and velocity distributions in the boundary layer.

Though no significant difference from the author's analysis was attempted, the results may be of interest. Profiles which would reduce in the time-dependent one-dimensional case to the exact solution, at least for Pr = 1, were chosen. This, of course, precludes the possibility of a finite boundary layer.

Analogous to Equations [3] and [4] of the paper

$$\frac{U}{U_1} = \eta\left[\frac{e^{-\eta^2}}{\sqrt{\pi}} - \eta \operatorname{erfc} \eta\right] \quad [86]$$

and

$$\frac{\theta}{\theta_w} = \operatorname{erfc} \eta \quad [87]$$

where

$$\eta = \frac{Y}{\Delta}$$

and

$\Delta = \Delta(x, T)$ = boundary-layer thickness parameter, ft.

A similar temperature profile and slightly different velocity profile were used recently for the two-dimensional steady state problem by Rutkowski.[14]

Upon integrating over the boundary layer (to $Y = \infty$) equations similar to Equations [5] and [6] of the paper are obtained but with slightly different constants

$$\frac{1}{6}\frac{\partial(u_1\delta)}{\partial\tau} + a\frac{\partial}{\partial x}(u_1^2\delta) = \mathrm{Gr}_L\mathrm{Pr}^2\delta - \mathrm{Pr}\frac{u_1}{\delta} \quad [88]$$

and

$$\frac{\partial}{\partial\tau}(\delta) + b\frac{\partial}{\partial x}(u_1\delta) = \frac{2}{\delta} \quad [89]$$

where

$$a = \frac{3\sqrt{2}}{20} - \frac{1}{5}$$

and

$$b = \frac{\sqrt{2} - 1}{6}$$

For the one-dimensional time-dependent case, δ and u_1 are not functions of x and solving Equations [88] and [89] of this discussion for Pr = 1 gives the exact solution of Illingworth,[15] and Equations [29] and [30] of the paper.

[13] Instructor, Mechanical Engineering Department, University of Minnesota, Minneapolis, Minnesota. Mem. ASME.

[14] ASME Paper No. 57—S-7.

[15] Refer to author's Bibliography (1).

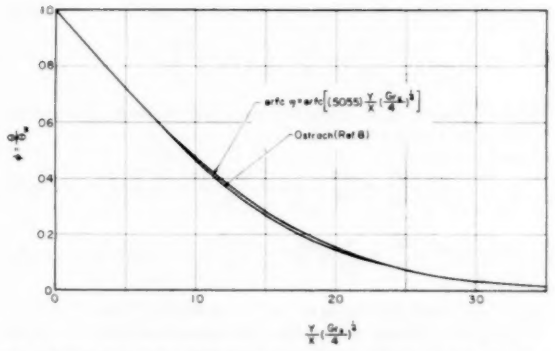


FIG. 12 STEADY-STATE TEMPERATURE DISTRIBUTION FOR PR = 1

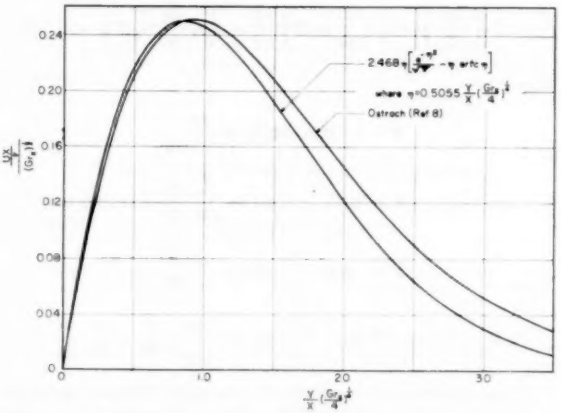


FIG. 13 STEADY-STATE VELOCITY DISTRIBUTION FOR PR = 1

For the steady-state case u_1 and δ are no longer functions of time and

$$\frac{\Delta}{X} = \mathrm{Gr}_x^{-1/4}\mathrm{Pr}^{-1/2}\left(\frac{8}{3b}\right)^{1/4}\left[\frac{10a}{3b} + \mathrm{Pr}\right]^{1/4} \quad [90]$$

and

$$\frac{U_1X}{\alpha} = \mathrm{Gr}_x^{1/2}\mathrm{Pr}\left(\frac{8}{3b}\right)^{1/2}\left[\frac{10a}{3b} + \mathrm{Pr}\right]^{-1/2} \quad [91]$$

Comparing the solutions (with Pr = 1) for velocity and temperature to the exact solutions of Illingworth and Ostrach,[16] both the temperature and velocity fields are exact in the transient one-dimensional case and fairly good at least close to the wall in the steady state[17] (see Figs. 12 and 13, herewith). In fact the heat-transfer coefficient at the wall is only about 0.5 per cent different from that determined by Ostrach.

Of greater interest is the length of time during which portions of the plate lose heat solely by conduction as this may show whether there is a minimum in the heat-transfer coefficient. Equations [88] and [89] of this discussion are hyperbolic and, following the author's analysis, the time for the end of pure conduction is

$$T = 2.37(\mathrm{Pr} + 1)^{1/3}(g\beta\theta_w)^{-1/3}X^{1/3} \quad [92]$$

[16] Refer to author's Bibliography (8).

[17] From the momentum and energy differential equations one can obtain the boundary conditions at the wall $\partial^3\theta/\partial Y^3 = 0$ and $\nu(\partial^3U/\partial Y^3) = -g\beta\theta$. In the steady state the first of these is met by the assumed profiles [86] and [87], but not the latter.

while the time at which steady state is reached is

$$T = 5.72\,(\mathrm{Pr} + 0.586)^{1/6}(g\beta\theta_\infty)^{-1/2}X^{1/3} \quad \ldots\ldots [93]$$

These are also quite similar to the author's results with only some of the constants being different.

When the heat-transfer coefficient is calculated not only is the steady-state value greater than the value at the end of one-dimensional conduction but, for Pr = 1, it is 30 per cent greater using these profiles. This compares with 26 per cent found from the author's profiles.

It should be borne in mind that although the Prandtl number is included in the analysis, the ability of the profiles to match both the transient and steady-state situation has been shown to be approximately valid only for Pr = 1. A large departure from Pr = 1 might give quite different results.

E. M. Sparrow.[18] The author is to be complimented on a well-executed and clearly presented analytical study. This paper is a most welcome addition to the little investigated field of transient convective heat transfer.

An especially interesting and somewhat surprising finding is that during the latter part of the transient period, the heat-transfer coefficient is lower than the steady-state value. Qualitative support of this result may be found in a recently published experiment by Ostroumov.[19] The apparatus consisted of a fine platinum wire, 0.1 mm diam and 107 mm long, stretched horizontally in a tank which was filled with either ethyl alcohol, water, or air. The thermal response of the wire to a suddenly applied direct current was studied. For the wire in alcohol, a plot of wire temperature as a function of time displayed a maximum in the latter part of the transient period which exceeded the steady-state temperature. A similar, but less pronounced overshoot was noted for the wire immersed in water. However, there was no observed overshoot in the air tests. Ostroumov offers no explanation for the different behavior encountered in the air tests, and it would appear that further experiments are needed. But, the fact that overshoot did occur in the experiments seems to lend support to the author's analysis.

Author's Closure

The author would like to thank the discussers for their interesting comments which supplement the content of the paper. Mr. Goldstein's analysis shows that a minimum in the transient heat-transfer coefficient is still obtained when another choice of velocity and temperature profiles is made for the isothermal plate. His profiles, chosen such that the integral method would yield the exact solution during the initial transient period dependent on time only, were also found to produce good results at steady state. The experiment mentioned by Dr. Sparrow indicates that this heat-transfer minimum may actually exist for free convection in some fluids.

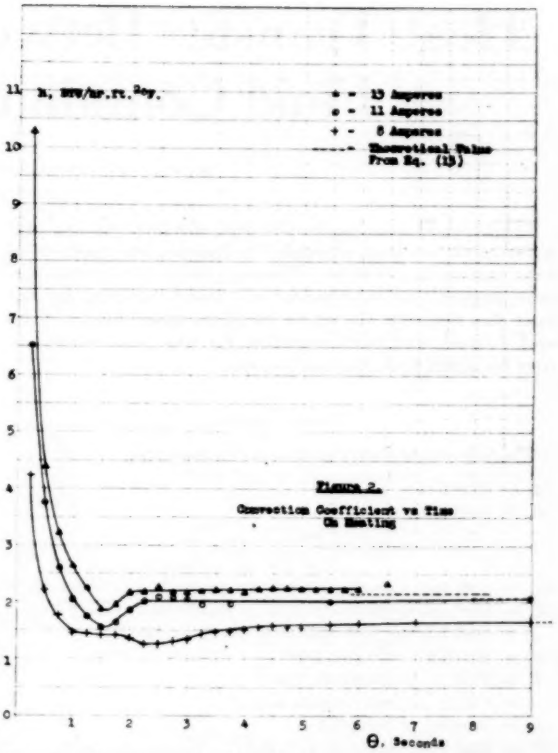


Fig. 14 Experimental Time Variation in Heat Transfer Coefficient (From H. Klei.)

Some additional experimental information from a thesis[20] by H. Klei was recently brought to the author's attention.[21] In this work a vertical metallic foil was heated stepwise electrically and the transient heat transfer to the surrounding air was measured. The test section was the central 1.5-in. length of a platinum strip 18 in. long, 1 in. high, and 0.0005 in. thick. The heating was such that the maximum temperature differences attained at steady state for three experimental runs were 199, 174, and 91 F above air at 70 F and one atmosphere pressure. The experimental results for heat-transfer coefficient as a function of time are shown on Fig. 14 (fig. 2 of Klei's thesis). They show that h does, in fact, display a minimum during the transient as predicted by the theoretical analysis.

[18] Heat Transfer Branch, NACA, Lewis Flight Propulsion Laboratory, Cleveland, Ohio. Assoc. Mem. ASME.

[19] "Unsteady Thermal Convection About a Horizontal Cylinder," by G. A. Ostroumov, *Zhurnal Tekhnicheskoi Fiziki* (Russian), vol. 26, 1956, pp. 2720–2730.

[20] Herbert Klei, SB thesis in Chem. Eng., M.I.T., May, 1957.

[21] Personal communication from Prof. G. C. Williams, M.I.T.

Heat Transfer Between a Flat Plate and a Fluid Containing Heat Sources

By I. R. WHITEMAN,[1] LOS ANGELES, CALIF.

The Leveque solution for the case of a fluid flowing past a flat plate has been expanded to include the presence of heat sources in the fluid. Through the use of certain approximations, an expression has been obtained for the heat flux through the plate for given plate temperature and source, and an expression for the plate temperature for given heat flux and source.

Nomenclature

The following nomenclature is used in the paper:

A = constant
B = constant
b = velocity gradient $\left(\frac{\partial u}{\partial y}\right)_{y=0}$, 1/hr
C_p = unit heat capacity at constant pressure, Btu/slugs-deg F
K = thermal conductivity of fluid, Btu/hr-sq ft (deg F/ft)
q = heat flux per unit wall area, Btu/hr-sq ft
S = transform variable, 1/ft^2
T = $T(X, y)$ temperature, deg F
t = $t(S, y)$ Laplace transform of T
u = velocity of fluid, fph
W = fluid heat source, Btu/cu ft-hr
x = distance along plate, ft
X = $\frac{K}{b\rho c_p} x$, cu ft
y = distance normal to plate surface, ft
ρ = fluid density, slugs/cu ft
η = dummy variable

Introduction

The Leveque solution[2] is the "asymptotic" solution to the Graetz problem.

We wish to find the asymptotic solution when there are sources of heat present in the fluid.

From the energy equation

$$u\rho C_p \frac{\partial T}{\partial x} = K \frac{\partial^2 T}{\partial y^2} + W \quad \ldots\ldots [1]$$

we may rewrite, expressing the fluid velocity as $u = by$, and making the suitable substitution in the following form

$$\frac{\partial T}{\partial X} = \frac{1}{y} \frac{\partial^2 T}{\partial y^2} + \frac{W}{Ky} \quad \ldots\ldots [2]$$

The boundary conditions are

$$T(X, 0) = T_p$$

$$\lim_{y \to \infty} T(X, y) = 0$$

The second boundary condition is based on the premise that, with large y, there is no influence of the wall and subsequently no heat transfer.

Thus Equation [1] reduces to

$$u\rho C_p \frac{\partial T}{\partial x} = W$$

Since $u = by$, with large y, u becomes large and thus $\partial T/\partial X$ becomes vanishingly small.

Let us take the Laplace transformation with respect to X and rewrite as follows

$$\frac{\partial^2 t}{\partial y^2} - Syt = -\frac{W}{KS} \quad \ldots\ldots [3]$$

Solving for the complementary solution, we obtain a modified Bessel function of the first kind of order 1/3

$$t = A(S) y^{1/2} I_{-1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) + B(S) y^{1/2} I_{1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) \quad \ldots\ldots [4]$$

And solving for the particular solution by the method of "variation of parameters,"[3] we obtain

$$t = \left(-\frac{1}{3}\right)! \left(\frac{1}{3}\right)! \frac{W}{KS} \left\{ y^{1/2} I_{-1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) \int y^{1/2} I_{1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) dy - y^{1/2} I_{1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) \int y^{1/2} I_{-1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) dy \right\} \quad \ldots\ldots [5]$$

Evaluating the constants, we find from the first boundary condition that all of the solution goes to zero, except the first part of the complementary solution, and so

$$A(S) = \left(-\frac{1}{3}\right)! \left(\frac{1}{3} S^{1/2}\right)^{1/3} t_p \quad \ldots\ldots [6]$$

To evaluate the remaining constant, let us turn our attention to the contribution of the particular solution, Equation [5].

As $y \to \infty$, the following approximation holds

$$y^{1/2} I_{\pm 1/3}\left(\frac{2}{3} S^{1/2} y^{3/2}\right) \sim y^{1/2} \frac{e^{\left(\frac{2}{3} S^{1/2} y^{3/2}\right)}}{\sqrt{\left[2\pi\left(\frac{2}{3} S^{1/2} y^{3/2}\right)\right]}} \quad \ldots\ldots [7]$$

---

[1] Assistant Research Engineer, Engineering Department, University of California at Los Angeles.

[2] "Heat Transfer Notes," by L. M. K. Boelter, et al., University of California Press, Los Angeles, Calif., 1946, p. X-38.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 20, 1956. Paper No. 57—SA-4.

[3] "Advanced Calculus for Engineers," by F. B. Hildebrand, Prentice-Hall, Inc., New York, N. Y., 1949.

and the particular solution, Equation [5], takes the form

$$t = -\left(-\frac{1}{3}\right)!\left(\frac{1}{3}\right)!\frac{W}{KS}\left\{\int\left[I_{-1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right) - I_{1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right)\right]y^{1/3}\,dy\right\} y^{1/2}\frac{e^{\left(\frac{2}{3}S^{1/2}y^{3/2}\right)}}{\sqrt{\left[2\pi\left(\frac{2}{3}S^{1/2}y^{3/2}\right)\right]}} \quad [8]$$

The integral portion of Equation [8] may be written in terms of the modified Bessel function of the second kind,[3] and reduces to

$$\frac{\sqrt{3}}{\pi}\int K_{1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right)y^{1/3}\,dy \quad [9]$$

Before evaluating this integral, let us make the substitution

$$\xi = \frac{2}{3}S^{1/2}y^{3/2} \quad [10]$$

and the integral takes the form

$$\frac{\sqrt{3}}{\pi S^{1/2}}\int K_{1/3}(\xi)d\xi \quad [11]$$

For large values of ξ, the modified function has the asymptotic behavior[3]

$$K_{1/3}(\xi) \sim \sqrt{\frac{\pi}{2}}\,e^{-\xi}\xi^{-1/3} \quad [12]$$

The integral expression Equation [9] takes the form of a gamma function and has the value

$$\sqrt{\left(\frac{3}{2S}\right)}$$

Our particular solution now takes the form

$$t = -\left(-\frac{1}{3}\right)!\left(\frac{1}{3}\right)!\frac{W}{KS}\left\{\sqrt{\left(\frac{3}{2S}\right)}\right\} y^{1/2}\frac{e^{\left(\frac{2}{3}S^{1/2}y^{3/2}\right)}}{\sqrt{\left[2\pi\left(\frac{2}{3}S^{1/2}y^{3/2}\right)\right]}} \quad [13]$$

The entire solution may now be written utilizing the asymptotic expressions

$$t = \left[A(S) + B(S) - \left(-\frac{1}{3}\right)!\left(\frac{1}{3}\right)!\sqrt{\frac{3}{2}}\frac{W}{KS^{3/2}}\right] y^{1/2}\frac{e^{\left(\frac{2}{3}S^{1/2}y^{3/2}\right)}}{\sqrt{\left[2\pi\left(\frac{2}{3}S^{1/2}y^{3/2}\right)\right]}} \quad [14]$$

To satisfy the second boundary condition, the term in the square brackets must be zero and

$$B(S) = -A(s) + \left(-\frac{1}{3}\right)!\left(\frac{1}{3}\right)!\sqrt{\frac{3}{2}}\frac{W}{KS^{3/2}} \quad [15]$$

The solution may now be written in expanded form neglecting terms of higher order

$$t = \left(-\frac{1}{3}\right)!\left(\frac{1}{3}S^{1/2}\right)^{1/3}t_p y^{1/3}\left[I_{-1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right) - I_{1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right)\right] + \left(-\frac{1}{3}\right)!\left(\frac{1}{3}\right)!\sqrt{\frac{3}{2}}\frac{W}{KS^{3/2}}y^{1/2}I_{1/3}\left(\frac{2}{3}S^{1/2}y^{3/2}\right) - \frac{W}{KS}\frac{y^2}{2} \quad [16]$$

The heat flux is determined from

$$q(S) = -K\left(\frac{dt}{dy}\right)_{y=0} \quad [17]$$

and we have

$$q(S) = \frac{\left(-\frac{1}{3}\right)!KS^{1/3}t_p}{3^{2/3}\left(\frac{1}{3}\right)!} - \frac{3^{1/3}\left(-\frac{1}{3}\right)!W}{2^{1/3}S^{2/3}} \quad [18]$$

Taking the inverse transform, we may now write the expression for the heat flux. This is to be interpreted as a Stieltjes-type integral[4]

$$q(X) = \frac{K}{3^{2/3}\left(\frac{1}{3}\right)!}\int_0^X \frac{dT_p}{d\eta}(X-\eta)^{-1/3}\,d\eta - \frac{3^{1/3}\left(-\frac{1}{3}\right)!X^{1/3}W}{2^{1/3}\left(\frac{1}{3}\right)!} \quad [19]$$

$$q(X) = 0.539K\int_0^X \frac{dT_p}{d\eta}(X-\eta)^{-1/3}\,d\eta - 1.29X^{1/3}W$$

The expression for the temperature at the plate is

$$t_p(S) = \frac{3^{2/3}\left(\frac{1}{3}\right)!S^{-1/3}q(s)}{K\left(-\frac{1}{3}\right)!} + \frac{3^{2/3}\left(\frac{1}{3}\right)!}{2^{1/2}S^{2/3}}\frac{W}{K} \quad [20]$$

And taking the inverse transform, we have the expression for the temperature at the plate

$$T_p(X) = \frac{3^{2/3}\left(\frac{1}{3}\right)!}{\left(-\frac{2}{3}\right)!\left(-\frac{1}{3}\right)!K}\int_0^X (X-\eta)^{-2/3}q(\eta)\,d\eta + \frac{3^{2/3}\left(\frac{1}{3}\right)!X^{2/3}}{2^{1/3}\left(\frac{2}{3}\right)!}\frac{W}{K} \quad [21]$$

$$T_p(X) = \frac{0.511}{K}\int_0^X (X-\eta)^{-2/3}q(\eta)d\eta + 1.75X^{2/3}\frac{W}{K}$$

Application

Let us consider the case in which a fluid with heat sources flows over a plate at constant temperature T_p.

[4] "Forced Convection From Nonisothermal Surfaces," by M. Tribus and J. Klein, Heat Transfer—a Symposium held at the University of Michigan during the Summer of 1952, Engineering Research Institute, University of Michigan, Ann Arbor, Mich., 1953, pp. 211–235.

If we replace the free-stream temperature by the adiabatic wall temperature, then the problem with heat sources can be reduced to one with no sources.[5] The adiabatic wall temperature T_{aw} is obtained for the case of zero heat flux.

Thus from Equation [19] we may write

$$q(X) = 0.539\,K(T_p - T_0)X^{-1/3} - 1.29WX^{1/3} \quad [22]$$

and

$$q(X) = 0.539K(T_p - T_{aw})X^{-1/3} \quad [23]$$

Equating Equations [22] and [23] we have

$$T_{aw} - T_0 = 2.4\,\frac{W}{K}\,X^{1/3} \quad [24]$$

Fig. 1 shows the variation of adiabatic wall temperature with distance along the plate.

The heat flux may then be calculated from

$$q = h_c A(T_p - T_{aw}) \quad [25]$$

in which h_c is the conductance evaluated when $w = 0$.

[5] "Frictional Heating of Nonisothermal Walls," by Myron Tribus and J. E. Mahlmeister, Readers' Forum, *Journal of the Aeronautical Sciences*, vol. 22, 1955, p. 726.

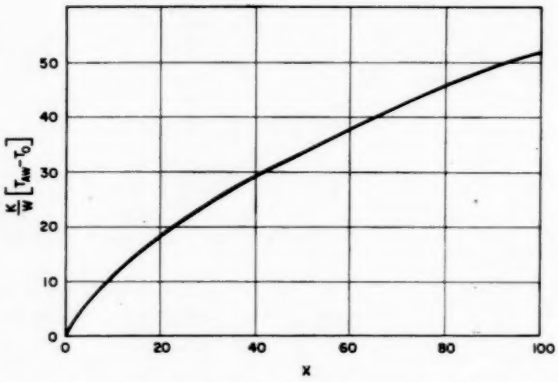


FIG. 1 VARIATION OF ADIABATIC WALL TEMPERATURE WITH DISTANCE ALONG THE PLATE

ACKNOWLEDGMENT

Thanks and appreciation must be extended to Dr. Myron Tribus of the University of California at Los Angeles for the suggestion of this problem and his *joie de vivre* of things scientific.

On the Stagnation of Natural-Convection Flows in Closed-End Tubes

By SIMON OSTRACH[1] and P. R. THORNTON,[2] CLEVELAND, OHIO

An analysis of the laminar natural-convection flow and heat transfer in a closed-end tube with a linear wall temperature and large but finite length-radius ratio is presented. It is found that for a given relation between the two physical parameters of the problem, the flow will fill the entire tube length. Representative velocity and temperature profiles are presented to show the effects of the parameters on the flow and heat transfer.

Introduction

THE application of natural-convection flows generated by large centrifugal forces for cooling rotating machinery has been of interest for several years. However, despite the numerous theoretical and experimental studies which recently have been reported on this subject in the literature, relatively little information on the actual flow and heat transfer in enclosed regions exists. One of the most interesting and, at the same time, distressing results that has been encountered, however, is that for a closed-end region with a sufficiently large length-diameter (or length-radius) ratio part of the fluid stagnates and, thus, is no longer effective as a coolant. This situation is predicted in reports[3,4] which treat the fully developed (i.e., infinite length-diameter ratio) natural-convection flow between two vertical plates with constant and linear wall temperatures, respectively, and by Lighthill[5] for a closed-end tube with large but finite length-radius ratio and constant wall temperature.

Since in actual configurations (see, for example, a paper by Schmidt[6]) at least one end of the coolant passage is closed, it is clear that additional information on this phenomenon would be of interest. Consideration is, therefore, given herein to the natural convection in a closed-end tube, but now the temperature will be taken to vary axially along the tube wall, Fig. 1. This generalization of the wall-temperature condition may be more realistic since the temperature in a turbine blade varies in the spanwise direction. The problem is solved as in the previous paper[3] by an integral method for large but finite length-radius ratios.

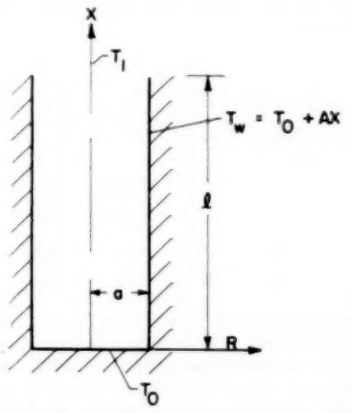


Fig. 1 Schematic Sketch of Configuration

Analysis

Basic Equations. The equations that will be used herein expressing the conservation of mass, momentum, and energy for steady axisymmetrical flow are, respectively

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial R} + \frac{V}{R} = 0 \quad [1]$$

$$U\frac{\partial U}{\partial X} + V\frac{\partial U}{\partial R} = \beta f(T - T_w) + \nu\left[\frac{\partial^2 U}{\partial R^2} + \frac{1}{R}\frac{\partial U}{\partial R}\right]_a^R \quad [2]$$

$$\frac{\partial P}{\partial R} = 0 \quad [3]$$

$$U\frac{\partial T}{\partial X} + V\frac{\partial T}{\partial R} = \kappa\left(\frac{\partial^2 T}{\partial R^2} + \frac{1}{R}\frac{\partial T}{\partial R}\right) \quad [4]$$

where X is measured axially along the tube from an origin at the closed end and R radially outward from the axis to the wall, Fig. 1; the corresponding velocities are U and V. The temperature is denoted by T, κ is the thermal diffusivity, ν the kinematic viscosity, β the volumetric-expansion coefficient, f the axial-body force component per unit mass, and P the pressure. The subscript w denotes wall conditions ($R = a$). The equations are similar to those for free-convection boundary-layer flows neglecting dissipation, except that the pressure no longer takes on its hydrostatic value. Hence, the wall is used as the reference condition for the buoyancy term as can be seen in Equation [2]. Justification for the use of the boundary-layer equations for all length-diameter ratios and details on the derivation of the buoyancy term are given in a previous report.[3]

[1] Chief, Applied Mechanics Branch, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

[2] Aeronautical Research Scientist, Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

[3] "Laminar Natural-Convection Flow and Heat Transfer of Fluids With and Without Heat Sources in Channels With Constant Wall Temperature," by S. Ostrach, NACA TN 2863, 1952.

[4] "Combined Natural and Forced-Convection Laminar Flow and Heat Transfer of Fluids With and Without Heat Sources in Channels With Linearly Varying Wall Temperatures," by S. Ostrach, NACA TN 3141, 1954.

[5] "Theoretical Considerations on Free Convection in Tubes," by M. J. Lighthill, *Quarterly Journal of Mechanics and Applied Mathematics*, vol. 6, 1953, pp. 398–439.

[6] "Heat Transmission by Natural Convection at High Centrifugal Acceleration in Water-Cooled Gas-Turbine Blades," by E. H. W. Schmidt, Proceedings of the General Discussion on Heat Transfer, The Institution of Mechanical Engineers and ASME, London, England, September, 1951, pp. 361–363.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 23, 1956. Paper No. 57—SA-2.

Equations [1] to [4] are nondimensionalized by letting

$$U = \frac{\kappa l}{a^2}\,u,\ V = \frac{\kappa}{a}\,v,$$

$$T = T_0 - \frac{\nu\kappa l}{\beta f a^4}\,t,\ X = lx,\ R = ar \quad \text{[5]}$$

where T_0 is the temperature along the closed end of the tube, l is the length, and a the radius of the tube. Thus, Equations [1], [2], and [4] become

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial r} + \frac{v}{r} = 0 \quad \text{[6]}$$

$$\frac{1}{Pr}\left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial r}\right) = -(t - t_w) + \left[\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right]_r \quad \text{[7]}$$

$$u\frac{\partial t}{\partial x} + v\frac{\partial t}{\partial r} = \frac{\partial^2 t}{\partial r^2} + \frac{1}{r}\frac{\partial t}{\partial r} \quad \text{[8]}$$

where Pr is the Prandtl number and $t_w \equiv t(x, 1)$.

Boundary Conditions. The physical boundary conditions to be imposed in this problem require that the tube be impermeable and that there is no slip at the surfaces; that is

$$U(X, a) = V(X, a) = U(0, R) = V(0, R) = 0$$

The thermal boundary conditions are

$$T(0, R) = T_0,\quad T(X, a) = T_0 + AX,\quad T(l, 0) = T_1$$

where A is the axial-temperature gradient at the wall. Applying Equation [5] to these yields

$$u(x, 1) = v(x, 1) = u(0, r) = v(0, r) = 0 \quad \text{[9]}$$

$$t(0, r) = 0,\quad t(x, 1) = -\mathrm{Ra}_A x \quad \text{[10]}$$

$$t(1, 0) = \mathrm{Ra}\,\frac{a}{l} \quad \text{[10a]}$$

where $\mathrm{Ra} = \beta f a^4(T_0 - T_1)/\nu\kappa$ is the Rayleigh number and $\mathrm{Ra}_A = \beta f a^4 A/\nu\kappa$ is a modified Rayleigh number based on the wall axial-temperature gradient. It is thus evident that the more general surface thermal condition (second of Equation [10]) introduces another physical parameter, namely, Ra_A in addition to that given by Equation [10a] which also occurred in the previous report.[8]

Integrated Equations. Equations [6] to [10] define a rather formidable boundary-value problem. Accordingly, an integral method will be employed to obtain a solution that satisfies integrated forms of the differential equations and the equations themselves at the tube axis and walls; namely

$$\int_0^1 ru\,dr = 0 \quad \text{[11]}$$

$$\frac{1}{\mathrm{Pr}}\frac{\partial}{\partial x}\int_0^1 ru^2 dr = -\int_0^1 r(t - t_w)dr + \frac{1}{2}\left(\frac{\partial u}{\partial r} - \frac{\partial^2 u}{\partial r^2}\right)_{r=1} \quad \text{[12]}$$

$$\frac{\partial}{\partial x}\int_0^1 rut\,dr = \left(\frac{\partial t}{\partial r}\right)_{r=1} \quad \text{[13]}$$

and at the wall and axis

$$\left(\frac{\partial^2 t}{\partial r^2} + \frac{1}{r}\frac{\partial t}{\partial r}\right)_{r=1} = 0 \quad \text{[14]}$$

$$\left(u\frac{\partial t}{\partial x}\right)_{r=0} = \left(\frac{\partial^2 t}{\partial r^2} + \frac{1}{r}\frac{\partial t}{\partial r}\right)_{r=0} \quad \text{[15]}$$

$$\frac{1}{\mathrm{Pr}}\left(u\frac{\partial u}{\partial x}\right)_{r=0} = -(t - t_w)_{r=0} + \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r}\frac{\partial u}{\partial r}\right)_1^0 \quad \text{[16]}$$

The continuity equation has been omitted here since Equation [11] replaces it and it is sufficient to solve for u and t only.

Solutions. It was pointed out previously[8] that there are three types of flow regimes which can exist in a closed tube and which are determined by the parameter $\mathrm{Ra}(a/l)$; that is, essentially by the length-to-radius ratio. Of particular interest in this paper is the large but finite length-radius regime. In this case, the tendency of the boundary layer to thicken with distance from the closed end has disappeared. The velocity and temperature distributions are similar at each section of the tube, only their scale increasing as the open end of the tube is approached. It can be seen from Equation [12] that u and t must have the same variation with x, and that this variation must be linear can be seen from either Equation [13] or [16]. These conclusions could be determined as easily by examination of Equations [6] to [8].

If we let u and t be the product of x times a polynomial in r^2 because of the symmetry of the problem, Equations [9] to [11] and [14] to [16] are satisfied by

$$u = -4\beta x(1 - 6r^2 + 9r^4 - 4r^6) - \frac{x}{24}\left(\mathrm{Ra}\,\frac{a}{l} + \mathrm{Ra}_A + \frac{16\beta^2}{\mathrm{Pr}}\right)(r^2 - 3r^4 + 2r^6) \quad \text{[17]}$$

$$t = x\left\{\mathrm{Ra}\,\frac{a}{l} - \beta\,\mathrm{Ra}\,\frac{a}{l}\,r^2 + \left[\frac{8\beta\,\mathrm{Ra}\,\frac{a}{l} - 9\left(\mathrm{Ra}_A + \mathrm{Ra}\,\frac{a}{l}\right)}{5}\right]r^4 + \left[\frac{4\left(\mathrm{Ra}\,\frac{a}{l} + \mathrm{Ra}_A\right) - 3\beta\,\mathrm{Ra}\,\frac{a}{l}}{5}\right]r^6\right\} \quad \text{[18]}$$

The relationships among the parameters β, $\mathrm{Ra}(a/l)$, and Ra_A are then obtained from the two remaining Equations [12] and [13]

$$\mathrm{Ra}\,\frac{a}{l} + \mathrm{Ra}_A + \frac{16\beta^2}{\mathrm{Pr}} = 320\left\{\frac{126\left[1 + \frac{\mathrm{Ra}_A}{\mathrm{Ra}(a/l)}\right] - 9\beta\left[1 + \frac{\mathrm{Ra}_A}{\mathrm{Ra}(a/l)}\right] - 42\beta - 2\beta^2}{13\left[1 + \frac{\mathrm{Ra}_A}{\mathrm{Ra}(a/l)}\right] - \beta}\right\} \quad \text{[19]}$$

and

$$\frac{1}{\mathrm{Pr}}\left\{\frac{48}{35}\beta^2 + \frac{\beta\left[\frac{16\beta^2}{\mathrm{Pr}} + \mathrm{Ra}(a/l) + \mathrm{Ra}_A\right]}{420} + \frac{\left[\frac{16\beta^2}{\mathrm{Pr}} + \mathrm{Ra}(a/l) + \mathrm{Ra}_A\right]^2}{120{,}960}\right\} = \frac{24[\mathrm{Ra}(a/l) + \mathrm{Ra}_A] + 7\beta\,\mathrm{Ra}(a/l)}{120} - 48\beta + \frac{8\beta^2}{\mathrm{Pr}} \quad \text{[20]}$$

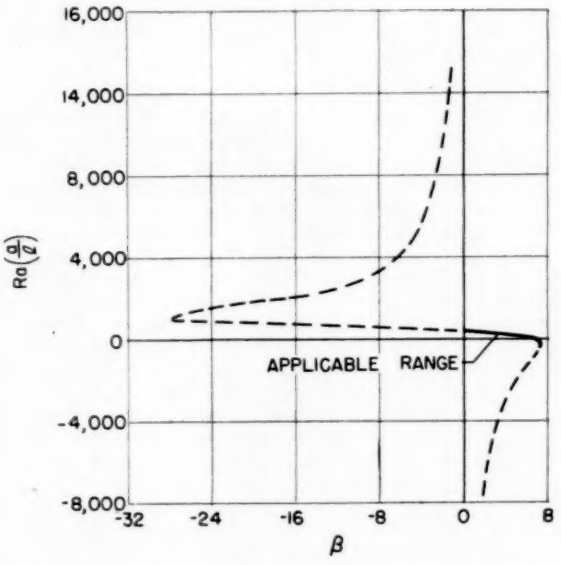


FIG. 2 RELATION BETWEEN THE PARAMETERS β AND $Ra(a/l)$ FROM EQUATION [21]

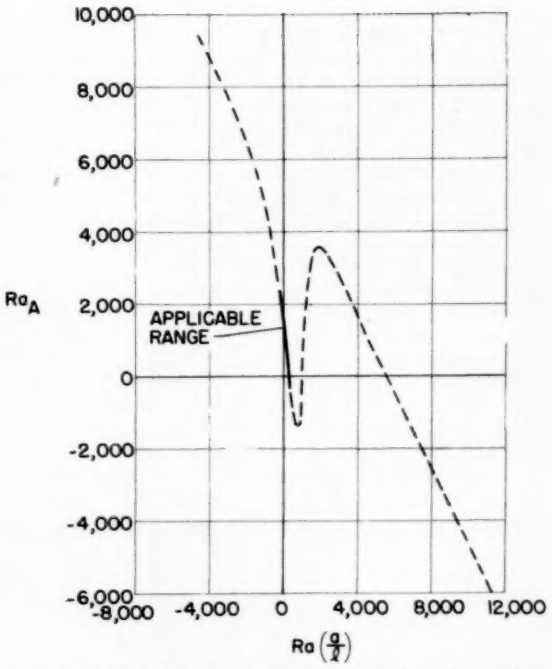


FIG. 3 RELATION BETWEEN THE PARAMETERS $Ra(a/l)$ AND Ra_A FROM EQUATION [20] FOR VALUES OF β FROM EQUATION [21]

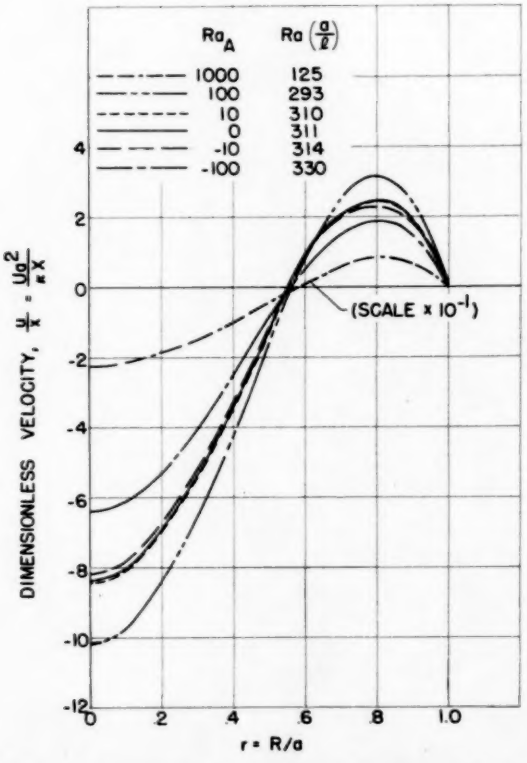


FIG. 4 REPRESENTATIVE VELOCITY DISTRIBUTIONS FOR VARIOUS VALUES OF Ra_A AND $Ra(a/l)$

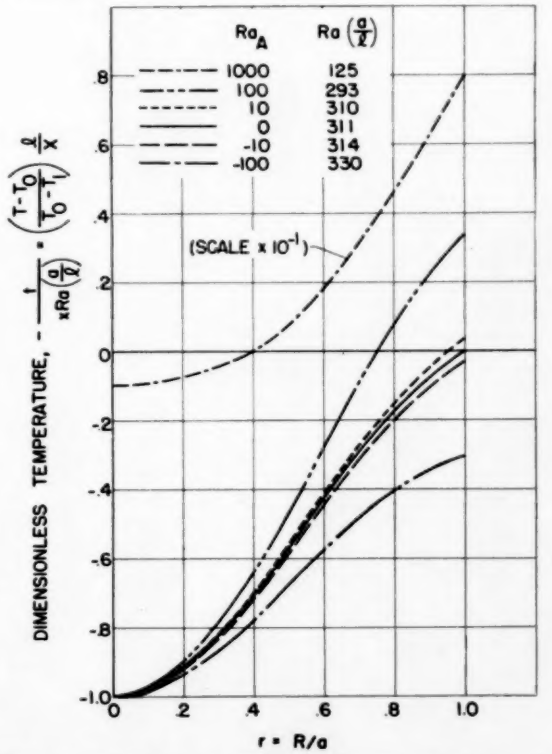


FIG. 5 REPRESENTATIVE TEMPERATURE DISTRIBUTIONS FOR VARIOUS VALUES OF Ra_A AND $Ra(a/l)$

For $Ra_A = 0$ Equations [19] and [20] reduce to Lighthill's[5] and, assuming $1/Pr = 0$, he then obtains three pairs of β, $Ra(a/l)$ values only one of which he argues has physical significance. (Justification for assuming $1/Pr = 0$ is discussed in his report.[5]) He thus determines a single value of $Ra(a/l)$ for which the solutions, Equations [17] and [18], hold. Therefore, the condition that the axial temperature should rise from its value T_1 at the open end of the tube to the value T_0 at the closed end determines the length-radius ratio for this flow. If the actual l/a is larger than this value, Lighthill points out that the excess length is filled with stagnated fluid.

If we note that Lighthill's problem[5] is a special case (with $A = 0$) of that considered herein it is reasonable to expect that the present problem will retain, in a sense, an eigenvalue character as discussed.[5] However, since there are two parameters in the present problems the solutions will be subject to a specific relation between them; that is, flows will be indicated for a range of parametric values rather than for discrete values (as in Lighthill's case) which implied the existence of the stagnation regions. Thus, for $Ra_A \neq 0$ as is postulated herein and also taking $1/Pr = 0$ there is obtained by eliminating Ra_A between Equations [19] and [20]

$$6{,}924{,}020.86 - 1{,}030{,}360.2\beta + 1{,}617.09\beta\, Ra(a/l) - 18{,}031.30\, [Ra(a/l)] - \beta[Ra(a/l)]^2 = 0 \quad [21]$$

This relation is plotted in Fig. 2. From this figure β can be determined for a given $Ra(a/l)$. For each value of β then Equation [20] with $1/Pr = 0$ yields a relation between $Ra(a/l)$ and Ra_A which is presented in Fig. 3.

Results

Velocity and temperature profiles can be determined with the use of Figs. 2 and 3. For a given Ra_A in the range $-1254 < Ra_A < 3579$ it is clear that three pairs of β, $Ra(a/l)$ values are obtained. Beyond this range the temperature gradients may be too large for similar flows. However, by arguments similar to those of Lighthill[5] only one pair for each Ra_A leads to physically meaningful profiles. The solid parts of the curves of Figs. 2 and 3 yield the reasonable profiles. Representative velocity and temperature distributions are presented in Figs. 4 and 5, respectively, and it can be seen that for each $Ra(a/l)$ there is an Ra_A for which a "similar" flow will exist in the entire tube. Lighthill's result[5] appears as a special case with $Ra_A = 0$. Of course, in an actual configuration it may not be possible to fix the axial-temperature gradient to the proper value in which case stagnation regions may occur. In any event, this simplified analysis does indicate that with proper design stagnation regions possibly could be eliminated in practical configurations.

The effects of the parameters on the velocity, temperature distribution and, hence, heat transfer can be seen in Figs. 4 and 5. Velocities and heat-transfer rates greater than those for the constant wall temperature ($Ra_A = 0$) case are obtained with positive values of Ra_A and the associated smaller values of $Ra(a/l)$.

A Model Method for Determining Geometric Factors in Solid-to-Solid Radiation Heat Transfer[1]

By P. L. TEA, JR.,[2] AND H. D. BAKER,[3] NEW YORK, N. Y.

A model method using light is presented for determining the geometric factors which must be known in order to utilize the Stefan-Boltzmann equation for heat transfer by solid-to-solid radiation. The model source is of unique design and closely approximates a uniform, perfectly diffuse plane source of any shape. The detector of radiation has negligible cosine error. The technique is highly suited to handle problems involving interreflections.

Nomenclature

The following nomenclature is used in the paper:

I_D = dark current from photomultiplier tube, amp
I = bucking current, amp
I_P = anode current of photomultiplier tube, exclusive of dark current, amp
V = bucking voltage
F_{ij} = configuration factor (also known as shape factor or angle factor), which is fraction of radiant energy emitted by area A_i which is directly incident on area A_j
ρ = simulated reflectivity of a receiving surface in the model
ρ_c = reflectivity of coating used on receiving surfaces
H = irradiancy, which is radiant energy incident per unit area per unit time

Introduction

In theory, a complete analysis of a heat-transfer problem by the mechanism of solid-to-solid radiation through a nonabsorbing medium presupposes a knowledge of all surface temperatures, emissivities of surfaces, degrees of diffuseness of all emitted energy, the natures of all reflections (diffuse, specular, or in between) of all emitting and receiving surfaces, and their geometrical relationships. Several ingenious methods of attack are available for determining the geometric factors needed in the Stefan-Boltzmann equation. For blackbody receivers, one may resort to exact or approximate mathematical means, descriptive geometry, mechanical or optical integrating devices, or models. If interreflections occur, we may add integral-equation theory, incremental methods, electric analogs, and digital-computer schemes.

Most methods postulate perfectly diffuse emission and reflection of radiation—that is to say, of the same angular distribution as blackbody radiation. Admittedly, not all problems fall within this realm. The model technique described herein is no exception; the attempt is made to approach conditions of perfectly diffuse emission and reflection, although the model could be adapted easily for conditions where reflections are other than perfectly diffuse.

Basically, the model, Fig. 1, comprises a visible-light source, a detector unit, and receiving surfaces, all of unique design. The aims are fourfold: (1) A working tool of moderate cost, with which a diversity of configurations may be set up and tested quickly and accurately; (2) simulation of a uniform, perfectly diffuse plane source of radiation of any shape; (3) a detector unit which is linear in its response, virtually immune to fatigue, and which is equally sensitive to radiation arriving at any angle of incidence, i.e., negligible cosine error; and (4) receiving surfaces which, when desired, can simulate perfectly diffuse reflection or reradiation, regardless of the angular distribution of the incident flux.

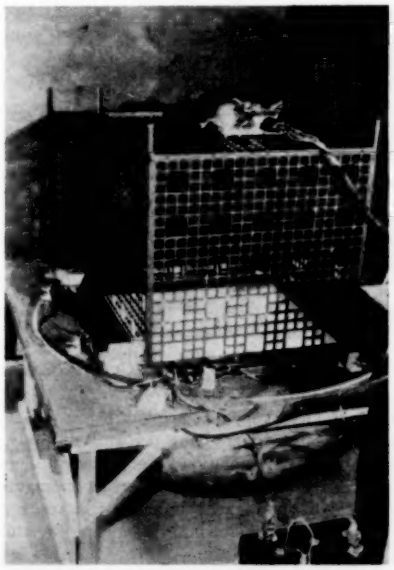


Fig. 1 Over-All View of Model With Cubical Chamber

The objectives of the tests to be described are to check the uniformity of the source, and to check calculated results for two model chambers. In one chamber, receivers are perfectly absorbing, and in the other, interreflections occur.

Theory and Construction of Source

Light, of roughly the visible range, serves as radiation. The

---

[1] This paper is based on a portion of the Doctoral research carried out by Peter L. Tea, Jr., at Columbia University in the Department of Mechanical Engineering, supported by the fellowship plan of E. I. du Pont de Nemours and Company and by the Eugene Higgens Trust.

[2] Department of Physics, The City College, College of the City of New York.

[3] Professor, Department of Mechanical Engineering, Columbia University. Mem. ASME.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 14, 1956. Paper No. 57—SA-10.

apparent source in the model is a 3-ft aluminum hemisphere, with its opening facing upward. The inner surface of the hemisphere was given three coats of white enamel followed by four coats of a special diffuse white paint, marketed by Benjamin Moore & Co. (New York, N. Y.) for coating integrating spheres used by photometry laboratories for measurements of mean spherical candlepower. Eight symmetrical spots around the hemisphere's periphery are illuminated directly by 3.2-watt miniature lamps powered by two 6-volt storage batteries. These lamps are enclosed by housings as indicated in Fig. 2, and housing slits limit the directly illuminated spots to rectangular areas 3 in. wide × 1/4 in. deep.

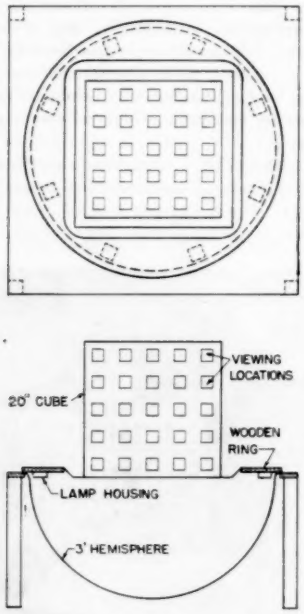


FIG. 2 SCHEMATIC MODEL WITH CUBICAL CHAMBER

According to the well-known principle of the integrating sphere, if the inside of a hollow sphere has a perfectly diffusely reflecting surface, the intensity of radiation reflected from all points due solely to interreflections within the sphere, as distinguished from reflection from spots illuminated directly by lamps, is equal and perfectly diffuse. For an incomplete sphere—in this case, a hemisphere—the principle still applies. Furthermore, when uniform, diffuse radiation is provided by the inside of a hemisphere either by self-emission, or as here by interreflections, the diametral plane of the hemisphere acts as a plane-disk source of the identical emissive power as the hemisphere source.[4]

The entire diametral plane of the hemisphere need not be utilized as the source in the model. Any portion of the plane may be used, and the remainder covered over, provided the surface which blocks off part of the plane is nonreflecting on its under side. Black velvet serves admirably. As may be seen from Fig. 2, the wooden ring, from which hang the lamp housings, obscures a portion of the diametral plane.

Precautions must be taken lest the radiation-detecting device "see" the directly illuminated spots when a reading is taken. Examination of Fig. 2 shows that this does not occur at any viewing location for the cubical chamber depicted.

[4] See, for example, "Lighting Design," by P. Moon and D. E. Spencer, Addison-Wesley Press, New York, N. Y., 1948, pp. 143–146.

It is interesting to observe that to the human eye the model source gives the illusion of being flat.

DETECTOR UNIT AND CIRCUIT

Referring to Fig. 3, a 3-in. hollow aluminum sphere (A) collects radiant flux through a 3/8-in-diam entrance aperture in a thin, flat annulus (B). A Radio Corporation of America (New York, N. Y.) 931-A photomultiplier tube (C) receives radiation reflected from a small area of the inner surface of the sphere through a second, much smaller hole (D). The defining slit (E) of the photomultiplier is formed by black tape on the glass wall of the tube. A rigid aluminum housing (F) supports the sphere and tube, and has four feet (G) to locate the entrance aperture of the sphere at

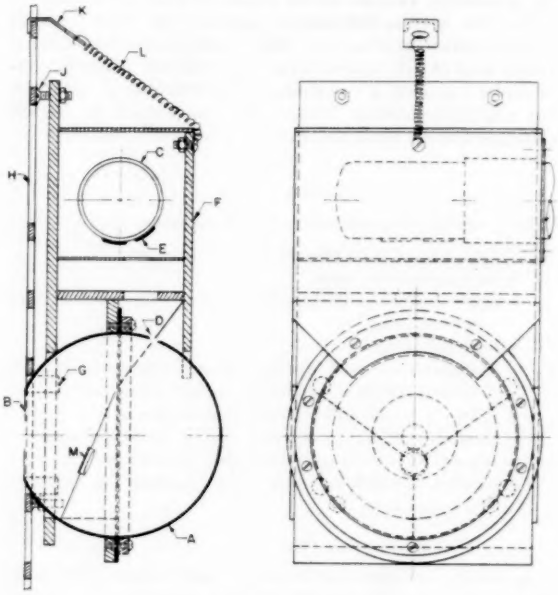


FIG. 3 DETECTOR UNIT

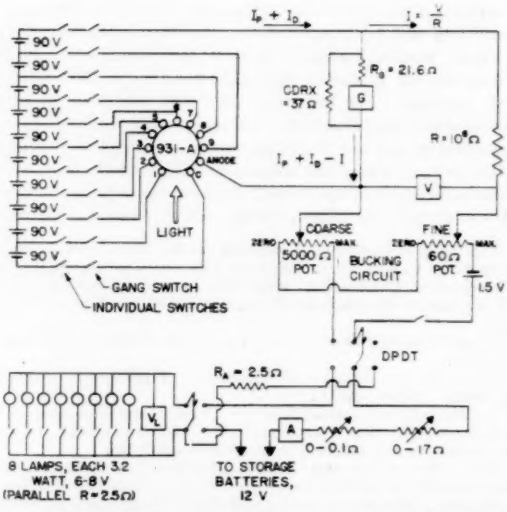


FIG. 4 CIRCUIT DIAGRAM FOR DETECTOR UNIT AND SOURCE LAMP

the center of any of the 1³/₄-in-square viewing locations which exist in a receiving surface (H). The plane of the entrance aperture is positioned in the plane of the inner surface of the chamber wall by means of four flat-head screws (J) which project from the housing and bear on the outer surface of the wall. A clamp (K), connected to the housing by a spring (L), hooks the unit in place.

Suspended within the sphere is a screen-and-lens assembly (M), which prevents the small area of the sphere wall which is "seen" by the photomultiplier from receiving full direct flux from the entrance aperture. The screen is perforated with 19 tiny, evenly spaced holes. Thus a definite fraction of any direct radiation incident on the screen is "transmitted" by these holes and is concentrated by the lens onto the area seen by the photomultiplier. The size of the holes in the screen was determined carefully by test, such that a reading for direct flux impinging on the screen will be identical with the reading which results from the same quantity of direct flux incident at any point of the sphere wall.

The collecting sphere is an integrating sphere, coated in exactly the same fashion as the hemisphere source. The inner surface of the annulus is painted a flat black.[5]

Long flexible leads join the photomultiplier with its power supply, Fig. 4. The anode current is opposed by an adjustable bucking circuit to give a null reading on a sensitive galvanometer (G) mounted on a Julius suspension, and used with a lamp-and-scale arrangement. Maximum anode currents employed were under 0.1 μ amp. For these conditions, anode current is reliably linear with illumination, and fatigue is slight. Bucking voltage V is directly proportional to the flux collected by the sphere, provided this flux is always of the same spectral distribution. Thus the circuit is capable of measuring relative—not absolute—values of irradiancy on the entrance aperture of the sphere.

Cosine error is close to zero, even for flux entering the sphere at large angles of incidence measured to the axis of the entrance aperture. This results from the thinness of the annulus, which is cut from 0.001-in. stainless-steel shim stock.

Experimental Procedure

Before beginning a run, source lamps and the photomultiplier were given a "warming-up" period of one hour. When a test was in progress, the laboratory, of course, was darkened completely except for the source in the model.

The first reading at a given receiving surface was made at what will be referred to as the principal viewing location. Every fifth reading was a repeat of the first. If the reading at the principal viewing location was no longer the same, the current to the source lamps was adjusted slightly to duplicate the original reading. Usually it took about ten minutes to complete a group of five readings, and the drift in the original reading seldom exceeded two per cent, and was usually much less. Corrections were applied to the four intervening readings, based on the assumption that the drift in sensitivity had proceeded at a uniform rate. Drift mostly tended to remain either positive or negative for several successive groups of readings. It is believed that the drift stemmed from three causes, either singly or in combination: (1) Gradual drop in emf of the storage batteries; (2) slight temperature changes, which affected resistances in the lamp circuit and the bucking circuit, and also influenced photomultiplier sensitivity; and (3) fatigue in the photomultiplier.

While a run was in progress, the photomultiplier was never switched off. Referring to Fig. 4, the lamp circuit and the bucking circuit could be switched on simultaneously by a double-pole, double-throw switch (DPDT). Before a reading was taken, with the detector unit in place, both circuits were switched off. The ground-glass scale was adjusted laterally to bring the hair-line image in the reflected spot of light from the galvanometer mirror in coincidence with zero on the scale. This adjustment, when needed at all, was necessitated by "zero drift" in the galvanometer suspension following a deflection. Under these initial conditions, only the dark current I_D from the photomultiplier, which had a steady value of about 2×10^{-3} μ amp flowed through the galvanometer. Meanwhile, the identical current which was later to flow to the source lamps was flowing through an auxiliary resistor R_A, so as not to afford the storage batteries a period of idleness during which a slight recovery in emf might ensue.

To take a reading, source lamps and bucking circuit were thrown in simultaneously. Coarse and fine controls of the bucking circuit were adjusted until the bucking current I equaled the anode current from the photomultiplier I_P, as indicated by a null reading. Thus

$$I_P = I = \frac{V}{R} = V \times 10^{-6} \quad [1]$$

The bucking voltage V was recorded.

All results were modified for a source of unit emissive power. Alternate readings were made at the principal viewing location and the center of the source. The average of ten readings at the principal viewing location divided by the average of ten readings at the source gave the irradiancy at the principal viewing location for a source of unit emissive power. The method of computing configuration factors from experimental data appears later, with the description of tests with a cubical chamber.

Test of Uniformity of Source

An auxiliary device, Fig. 5, was fabricated to suspend and position the detector unit for taking readings in the plane of the open floor of the model looking into the hemisphere source. Readings were taken at the centers of 25 imaginary squares comprising a 20-in-square source lying in the diametral plane of the hemisphere. Actually, readings were taken with the plane of the entrance aperture $^1/_2$ in. below the plane of the floor, to avoid flux reflected directly from the eight illuminated spots from entering the collecting sphere. The average values for two readings at each of the 25 viewing locations, expressed in bucking voltages corrected for drift, are indicated in Table 1 in the pattern of viewing locations for easy reference.

Maximum variation in readings was 1.6 per cent. The center,

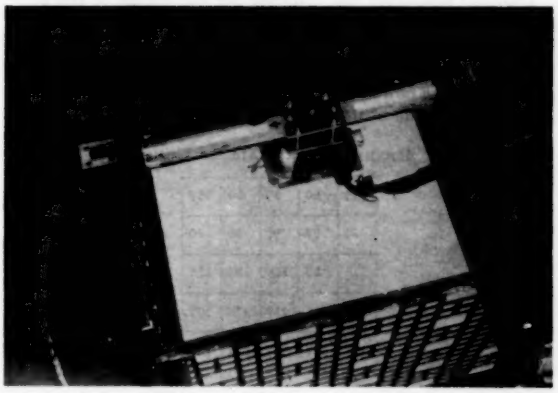


Fig. 5 Use of Auxiliary Device for Viewing Source

[5] "A New Detector Unit for Irradiance Measurements, Utilizing an Integrating Sphere and a Photomultiplier Tube," by P. L. Tea, Jr., and H. D. Baker, *Journal of the Optical Society of America*, vol. 46, 1956, pp. 875–878.

TABLE 1 READINGS FOR SQUARE SOURCE, $V \times 10^4$

866	873	873	872	867
878	867	866	868	875
871	868	866	867	877
876	868	864	865	874
866	873	871	873	864

which was the principal viewing location, was 0.4 per cent below average. In subsequent tests to be described, source readings were taken only at the center of the floor, and were accepted as being indicative of uniform emission from all points of the floor.

TESTS WITH A CUBICAL CHAMBER, BLACKBODY RECEIVERS

A 20-in. cube was assembled from sheets of aluminum $^3/_{32}$ in. thick, perforated with square holes as indicated in Figs. 1, 2, and 3 (left view). The surfaces of the cube were fastened by screws, so that it could be disassembled in a matter of minutes and the components used in the formation of other configurations. The open bottom of the cube served as source. All surfaces were painted a flat black. In addition, in order to simulate blackbody receivers closely, all surfaces except the one over which readings were being taken were lined inside with black velvet. Because of symmetry, readings were made only on one wall and the ceiling.

Wall readings are presented in Table 2, in the pattern of viewing locations. The ratio of the average of ten readings at the center viewing location of the bottom row of the wall to the average of ten readings at the center of the source was 0.421. It is significant that according to formula[6] the irradiancy at this wall location is calculated to be 0.423.

The sum of 25 wall readings, each representative of a 4-in. square, was 3839×10^{-4} volt. The configuration factor F_{fw}, from floor to wall was

$$F_{fw} = \left[\frac{\text{average wall reading}}{\text{reading at principal viewing location}}\right] \times \left[\frac{\text{average reading at principal viewing location}}{\text{average reading at center of floor}}\right] = \frac{0.3839}{25 \times 0.0327} \times 0.421 = 0.198 \quad [2]$$

By quadruple integration, the theoretical answer is known to be 0.200.

Identical procedure was followed for the ceiling of the cube. Results are shown in Table 3. Somewhat higher, source-lamp voltage was used, hence the higher readings as compared with the wall. Irradiancy at the center of the ceiling, based on a source of unit emissive power, was found to be 0.235, as compared with a theoretical value of 0.239. The configuration factor F_{fc}, from

TABLE 2 READINGS FOR WALL OF CUBE, $V \times 10^4$

55	60	69	60	57
80	97	96	96	80
116	127	145	129	116
160	215	217	214	171
267	312	327	305	267

[6] See "Radiant-Interchange Configuration Factors," by D. C. Hamilton and W. R. Morgan, NACA TN 2836, December, 1952.

TABLE 3 READINGS FOR CEILING OF CUBE, $V \times 10^4$

293	341	355	342	295
341	386	404	386	339
355	404	426	404	355
341	381	404	383	340
296	342	356	342	293

floor to ceiling, came out to be 0.197, as against the known theoretical value of 0.200 from a continuous uniform floor to a continuous ceiling of a cube.

TEST WITH A CYLINDER, WITH INTERREFLECTIONS

A right circular cylinder, $8^1/_4$ in. $\times$ $8^1/_4$ in., was positioned with its open lower end in the diametral plane of the hemisphere source, Fig. 6. The equivalent source was a nonreflecting radiating disk at the lower end of the cylinder, uniform and perfectly diffuse. The upper end of the cylinder was open, to simulate a blackbody receiver. Light passing through the open top was lost to the (black) walls of the laboratory.

To simulate desired reflectivities over various portions of the inner wall of the cylinder, groups of numerous, evenly spaced holes were punched in the cylinder. The coating of the wall that remained was identical with that of the hemisphere source, and it was assumed once again that it exhibited almost perfectly diffuse reflection. Flux passing through the holes was "absorbed." Simulated reflectivity for the wall ρ was taken to be

$$\rho = \rho_e \left[\frac{(\text{total area}) - (\text{combined area of holes})}{\text{total area}}\right] \quad [3]$$

where ρ_e was the reflectivity of the coating. In Fig. 6 is reproduced the curve of spectral reflectivity of the coating, as determined by test on a flat sample, performed by the Electrical Testing Laboratories, Inc. (New York, N. Y.), using a recording spectrophotometer. The curve is reasonably flat. After making a study of the relative response of the entire apparatus at the various wave lengths, it was decided that $\rho_e = 0.925$ was representative. Due consideration was given in this study to the esti-

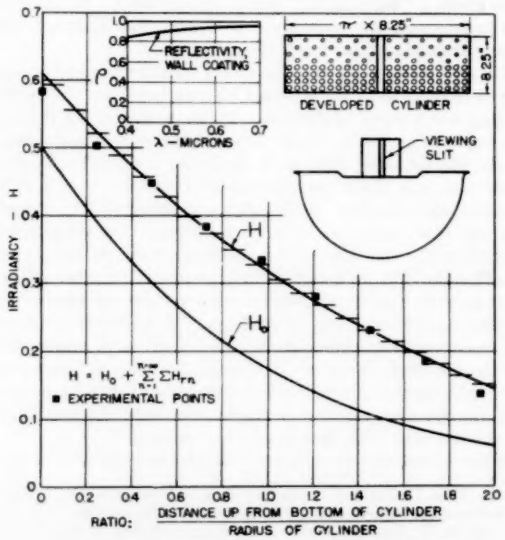


FIG. 6 CURVES FOR CYLINDER PROBLEM

mated spectral emission of the source lamps operating at 6.4 volts, to modifications in this spectral distribution created by interreflections in the hemisphere source and in the collecting sphere, and to the spectral response of the photomultiplier. Peak effect was apparently achieved by radiation at about 0.525 μ.

The developed cylinder is also shown in Fig. 6. For the lower half, the combined area of the holes represented 0.365 of the total area, and for the upper half of the cylinder, 0.127. Thus the simulated reflectivities for the lower and upper halves of the cylinder were, respectively, $0.925 \times (1.000 - 0.365) = 0.587$, and $0.925 \times (1.000 - 0.127) = 0.807$.

Readings were taken at nine viewing locations 1 in. apart along an element of the cylinder, through a 1-in. slit. Care was taken to screen the directly illuminated spots of the hemisphere source from the detector unit at its lowest position on the wall, which was at the base of the cylinder.

Alternate readings also were taken at the lowest wall position, and in the plane of the base of the cylinder looking into the source, ten at each position. The ratio of the average reading at the lowest wall position to the average looking into the source was 0.582. By multiplying all wall readings by an appropriate factor to refer them to a reading at the lowest wall position of 0.582, the irradiancy at each wall position was calculated for a source of unit emissive power. These experimental points are plotted in Fig. 6 and are found to be in good agreement with the plotted curve for irradiancy H which was calculated. Details of this calculation are presented in the Appendix.

The radiant energy absorbed per unit area at any point, for a source of unit emissive power, would be

$$H(1.000 - \rho) \quad [4]$$

Discussion

The excellent agreement between experimental results and theory for the cube and the cylinder, together with the good uniformity of the source, lends encouragement that this model technique may be used with confidence for other, more difficult configurations.

Unlike other techniques using light, in which a uniform source is constructed according to other principles, the source in this model is, in theory, not only uniform, but perfectly diffuse as well.

An important feature is that reflectivities over any or all receiving surfaces may be dissimilar; indeed, the reflectivity may differ over any surface, either continuously or discontinuously, by varying accordingly the size and spacing of absorbing holes.

The one-way heat transfer from the source to a given receiving surface is that portion of the emission from the source which is absorbed by that receiver, as determined by Equation [4]. Unless the receiver temperature is quite low this does not give the net radiant heat transfer. If, in the actual problem all surfaces are "gray," i.e., have emissivities which are independent of wave length, and all emissions and reflections are perfectly diffuse, a reciprocity relationship exists. Here, the percentage of emission from the source which is absorbed by a receiver is equal to the percentage of the (supposed) emission from the receiver which is absorbed by the source.[7] Where the receiver temperature is not uniform its area can be broken up into several "constant-temperature" zones; the number of such zones in the model technique is limited by the number of viewing locations on the receiver.

This model technique can be adapted for studies of illumination in rooms. No problem exists here concerning emission (of light) from receiving surfaces. Interest would be centered on illumination at various localities, not on absorption of light by the surfaces.

Use of the model for interreflection problems obviates the need for calculating in advance any configuration factors, either from the source to (blackbody) receivers, or between receivers.

[7] See "Heat Transmission," by W. H. McAdams, McGraw-Hill Book Company, Inc., New York, N. Y., third edition, 1954, pp. 72–76.

Acknowledgment

It is a pleasure to acknowledge the helpful suggestions given by Profs. L. J. Hayner and J. R. Roebuck.

Appendix

Incremental Solution of Cylinder Problem

The general technique for solving interreflection problems by increments was clearly enunciated by Moore.[8] The cylinder problem, described in the present paper and solved experimentally using the model technique, lends itself to an interesting solution by increments, using a desk calculator.

The notation is as follows:

H_0 = irradiancy at any point on inner wall of a blackbody cylinder of unit radius, caused by a uniform, diffusely emitting, nonreflecting disk of unit emissive power and unit radius at its base

y = distance from base of cylinder

E = apparent emissive power of an incremental cylinder, due to first reflection of direct flux from source

ρ = simulated reflectivity of cylinder wall

F_{ij} = configuration factor

A_1, A_2 = areas of incremental cylinders

Z_1, Z_2 = distances measured from an incremental cylinder

H_{rn} = irradiancy at one incremental cylinder from another incremental cylinder due to the nth reflection of flux from the source. The total contribution from all twenty incremental cylinders is ΣH_{rn}

H = total irradiancy, including interreflections, and

$$H = H_0 + \sum_{n=1}^{n=\infty} \Sigma H_{rn}$$

For the problem at hand, Hamilton and Morgan[6] give

$$H_0 = \frac{1}{2}\left[\frac{y^2 + 2}{(y^2 + 4)^{1/2}} - y\right] \quad [5]$$

The graph of H_0 versus y is shown in Fig. 6.

The cylinder was divided into 20 incremental cylinders. Cylinder No. 1 lay between $y = 0$ and $y = 0.1$, No. 2 between $y = 0.1$ and $y = 0.2$, and so forth. Referring to Fig. 7 (a), each incre-

[8] "Interreflections by the Increment Method as Applied to a Light Court," by A. D. Moore, Illuminating Engineering Society Transactions, vol. 24, 1929, pp. 629–670.

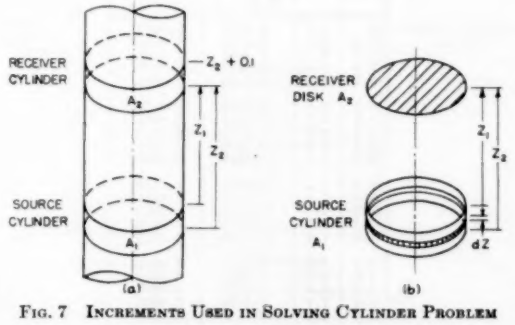


Fig. 7 Increments Used in Solving Cylinder Problem

mental cylinder was considered to be a source of apparent emissive power E due to the first reflection of direct flux from the source, where

$$E = \rho \text{ (average } H_0 \text{ on increment)} \quad [6]$$

Each of the 20 small cylinders received an addition to its irradiancy H_{r1}, from each of the 20 cylinders (including itself) as a result of the first reflection of H_0. The total contribution from all 20 cylinders for this first reflection was ΣH_{r1}.

The configuration factor F_{12} from any incremental cylinder A_1 acting as a source to any cylinder A_2 acting as receiver is the difference between configuration factors from A_1 to a disk at z_2, and from A_1 to a disk at $(z_2 + 0.1)$. It was first necessary to derive a formula for F_{12}', the configuration factor from A_1 to a disk at z_2. Referring to Fig. 7(*b*), and utilizing Equation [5], and recognizing that $F_{12}' = F_{21}'$

$$F_{12}' = F_{21}' = \frac{1}{(z_2 - z_1)} \frac{1}{2} \int_{z_1}^{z_2} \left[\frac{z^2 + 2}{(z^2 + 4)^{1/2}} - z \right] dz$$

$$= \frac{1}{4(z_2 - z_1)} \left[z_2(z_2^2 + 4)^{1/2} - z_1(z_1^2 + 4)^{1/2} - z_2^2 + z_1^2 \right] \quad [7]$$

Next, Table 4 was prepared. In the case of an incremental cylinder receiving radiation from itself

$$F_{11} = 1.000 - 2F_{12}' \quad [8]$$

where, in this case, F_{12}' is from the small cylinder to a disk at either end of itself.

It bears emphasis that each incremental cylinder in turn served as a source to each of the twenty incremental cylinders, and the z's were measured, in both directions, from the particular small cylinder source in question. The average irradiancy H_{r1} at any incremental cylinder resulting from the first diffuse reflection from any other cylinder was

$$H_{r1} = \frac{A_2 F_{12}}{A_1} E = \frac{2\pi(0.1)F_{12}}{2\pi(0.1)} \rho \text{ (average } H_0 \text{ on } A_1)$$

$$= F_{12}\, \rho \text{ (average } H_0 \text{ on } A_1) \quad [9]$$

In Table 5 is indicated the scheme by which ΣH_{r1} was obtained for each incremental cylinder. For incremental cylinders No. 1 through No. 10, $\rho = 0.587$, and for No. 11 through No. 20, $\rho = 0.807$. Next, the emission, by reflection, from each small cylinder due to the incidence of ΣH_{r1} was computed, and H_{r2} was computed for each cylinder in the same fashion. The procedure was repeated until the reflections had effectively died out. Tabulations for H_{r2}, H_{r3}, and so forth, are not given here for lack of space, but results appear in Table 6.

After ΣH_{r5} had been computed, reflections were quite weak, and a rough estimate was made of the sum of the subsequent ΣH_{rn}'s from $n = 6$ through $n = \infty$ based on the convergence of the series. The total irradiancy is listed in Table 6, and the curve appears in Fig. 6

$$H = H_0 + \sum_{n=1}^{n=\infty} \Sigma H_{rn} \quad [10]$$

Table 4 Calculation of F_{12}

I Receiver cylinder		II F_{12}' from cylinder between z = 0 and z = 0.1 to more distant disk for cylinder given in col. I	III F_{12} from cylinder between z = 0 and z = 0.1 to cylinder listed in col. I
From z =	To z =		
0	0.1	0.4756	0.0488
0.1	0.2	0.4294	0.0462
0.2	0.3	0.3868	0.0426
0.3	0.4	0.3478	0.0390
0.4	0.5	0.3124	0.0354
0.5	0.6	0.2802	0.0322
0.6	0.7	0.2511	0.0291
0.7	0.8	0.2250	0.0261
0.8	0.9	0.2015	0.0235
0.9	1.0	0.1806	0.0209
1.0	1.1	0.1618	0.0188
1.1	1.2	0.1452	0.0166
1.2	1.3	0.1303	0.0149
1.3	1.4	0.1171	0.0132
1.4	1.5	0.1054	0.0117
1.5	1.6	0.0950	0.0104
1.6	1.7	0.0858	0.0092
1.7	1.8	0.0775	0.0083
1.8	1.9	0.0702	0.0073
1.9	2.0	0.0637	0.0065

Table 5 Calculation of ΣH_{r1}

From cylinder No.	$H_{r1} \times 10^4$ on cylinder No.									
	1	2	3	4	5	6	7	8	9	10
1	137	129	118	109	99	90	81	73	66	58
2	117	123	116	107	98	90	81	73	66	59
3	96	105	111	105	96	89	80	73	66	59
4	80	87	94	100	94	87	80	73	66	59
5	65	72	78	85	90	85	78	72	65	59
6	53	58	64	70	76	81	76	70	64	58
7	43	48	52	57	63	68	71	68	63	57
8	34	38	43	47	52	56	61	65	61	56
9	28	30	34	38	42	46	51	54	58	54
10	22	25	28	30	34	37	41	45	49	52
11	24	28	30	34	38	42	46	51	55	60
12	19	22	25	28	30	34	38	42	46	50
13	15	17	20	22	25	27	30	33	37	41
14	12	14	15	18	20	22	25	28	30	33
15	10	11	14	14	16	18	20	22	25	28
16	8	9	10	11	14	14	16	18	20	22
17	7	7	9	10	10	11	13	15	16	18
18	5	6	7	8	9	10	10	11	13	14
19	4	5	5	6	7	8	9	10	10	11
20	4	4	4	5	6	6	7	8	9	10
ΣH_{r1}	783	838	877	904	919	921	914	904	885	858

From cylinder No.	$H_{r1} \times 10^4$ on cylinder No.									
	11	12	13	14	15	16	17	18	19	20
1	52	47	42	37	32	29	26	23	20	18
2	52	48	42	37	33	30	26	23	21	18
3	53	48	43	37	33	30	27	24	21	19
4	53	48	43	38	33	30	27	24	21	19
5	53	48	43	38	34	30	28	24	21	19
6	53	48	43	38	34	30	28	25	22	19
7	52	48	43	38	34	30	28	25	22	19
8	52	47	43	38	34	31	28	25	22	20
9	51	46	42	38	34	30	28	25	22	20
10	49	45	41	37	34	30	28	25	22	20
11	64	60	55	50	46	42	38	34	30	28
12	54	57	54	49	46	42	38	34	30	28
13	45	49	52	49	45	41	37	33	30	28
14	37	40	44	46	44	40	37	33	30	28
15	30	33	36	39	42	39	36	33	30	28
16	25	27	30	32	35	37	35	32	30	27
17	20	22	25	27	30	31	33	31	30	27
18	16	18	20	22	25	27	29	30	29	27
19	13	15	16	18	20	22	24	26	28	26
20	10	12	13	15	16	18	20	22	24	25
ΣH_{r1}	834	806	770	723	684	639	601	551	505	463

Table 6 Calculation of H

	$H \times 10^4$ on cylinder No.									
	1	2	3	4	5	6	7	8	9	10
H_0	4756	4294	3868	3478	3124	2802	2511	2250	2015	1806
ΣH_{r1}	783	838	877	904	919	921	914	904	885	858
ΣH_{r2}	243	264	284	300	315	329	338	345	352	356
ΣH_{r3}	91	100	108	115	123	127	134	137	145	145
ΣH_{r4}	36	40	43	47	50	52	55	58	60	60
ΣH_{r5}	13	13	17	20	21	23	24	24	25	26
ΣH_{rn}	10	10	14	16	16	17	18	18	19	19
H	5932	5559	5211	4880	4568	4271	3994	3736	3501	3270

	$H \times 10^4$ on cylinder No.									
	11	12	13	14	15	16	17	18	19	20
H_0	1618	1452	1303	1171	1054	950	858	775	702	637
ΣH_{r1}	834	806	770	723	684	639	601	551	505	463
ΣH_{r2}	356	353	348	337	328	311	293	278	257	236
ΣH_{r3}	148	147	147	144	141	137	131	123	115	105
ΣH_{r4}	62	61	63	63	60	59	56	53	49	45
ΣH_{r5}	26	27	26	26	24	24	24	22	20	17
ΣH_{rn}	19	19	19	19	18	18	18	17	16	14
H	3063	2865	2676	2483	2309	2138	1981	1819	1664	1517

Measurements of the Total Absorptivity for Solar Radiation of Several Engineering Materials[1]

By RICHARD C. BIRKEBAK[2] and J. P. HARTNETT,[3] MINNEAPOLIS, MINN.

Values are presented of the total solar absorptivity of several porous materials presently being considered for transpiration cooling of high-speed vehicles. To specify these surfaces photomicrographs and a chemical analysis are presented. Two schemes used in the measurement of the absorptivity values are described in the text, a comparison technique and an integrating radiometer method.

Nomenclature

The following nomenclature is used in the paper:

C = amount of energy reaching the thermopile from the test surface divided by the actual amount leaving the test surface (radiometer device)
e = amount of energy emitted by the surface, Btu/hr sq ft
F_{12} = shape factor, the amount of energy emitted per unit area from A_1 intercepted by A_2 divided by the amount emitted by A_1
G = radiation impinging on thermopile receiving surface, Btu/hr sq ft
J_s = incident solar energy, Btu/hr sq ft
k = proportionality constant
q_e = electrical energy added to surface when blocked off from sun to raise it to equilibrium temperature when exposed to solar energy, Btu/hr sq ft
q_l = heat lost by conduction and convection from surface, Btu/hr sq ft
T = absolute temperature R
α = total absorptivity
α' = total absorptivity for long wave radiation
ϵ = total emissivity
ρ = total hemispherical reflectivity
ρ' = reflectivity for long wave radiation
σ = Stefan-Boltzmann constant, 0.171×10^{-8} Btu/hr sq ft R^4
Δ = measured thermopile output

Subscripts

m = hemispherical mirrored surface
r = reference surface
s = test surface
t = thermopile
w = surrounding walls

[1] Publication of the Heat Transfer Laboratory, Mechanical Engineering Department, University of Minnesota.

[2] Instructor, Mechanical Engineering Department, University of Minnesota.

[3] Associate Professor, Mechanical Engineering Department, University of Minnesota. Assoc. Mem. ASME.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 4, 1957. Paper No. 57—SA-27.

Introduction

The knowledge of the total absorptivity for solar radiation (0.3–3.0 microns) of engineering materials is of importance in many applications. For example, when exposed to the sun, the equilibrium temperature of an airplane or missile is dependent on the balance between convective and radiative heat transfers and consequently the knowledge of the absorptivity for solar radiation is necessary to predict the temperature. Such measurements for materials of current interest in aircraft and missile applications have been obtained in the Heat Transfer Laboratory of the University of Minnesota utilizing two different experimental procedures. Of particular interest among the materials investigated are the results for porous materials presently being considered for transpiration cooling in high-speed applications.

Description of Apparatus

Two measurement techniques were used to determine the total absorptivity, (*a*) a comparison method, and (*b*) an absolute integrating radiometer method.

Absorptivity Comparison Instrument. This device which represents a modification of methods previously proposed by Dunkle and Gier (4)[4] and Wilkes (7) consists of a standard surface

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

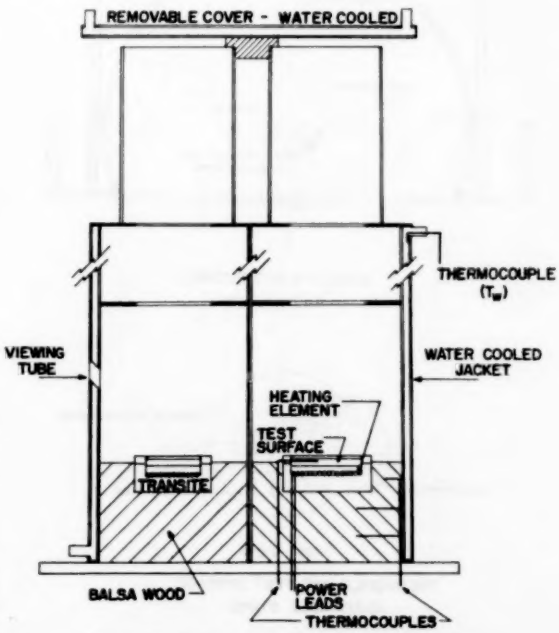


Fig. 1 Absorptivity Comparison Instrument

of known absorptivity and a test surface with unknown absorptivity, but of similar geometry. Fig. 1 shows the construction of the apparatus and the positions of the two surfaces. When exposed to the sun, the cover is removed and the solar energy is directed onto the reference and test surfaces, which are enclosed by water-cooled cylindrical chambers with blackened inner surfaces to minimize any stray reflections and to provide a constant immediate environment for the surfaces. To ensure that the sun's energy is always on the test areas, a viewing tube is available. A cover so designed as to close off the two surfaces completely from solar energy without inhibiting free convection to the surroundings was utilized during the test program.

Two heating elements, one for the reference surface, and one for the test surface are recessed in the holder, Fig. 1. The reference surface is coated with acetylene soot applied with a direct flame and has a solar absorptivity of 0.99 (3).

The whole assembly is mounted on a turntable to allow for following the sun throughout the test program.

Absolute Integrating Radiometer. The absolute integrating radiometer used in the second series of tests is shown in Fig. 2 and represents a modification of instruments used by Beckett (1) and Coblentz (3). A thermopile and a test surface are located on conjugate foci on the diameter of the hemisphere. If solar energy passes through the aperture directly onto the receiving surface of the thermopile, the reading of the galvanometer connected to the thermopile is an indication of the amount of incident solar energy. If the solar energy is next reflected off the test surface and onto the thermopile, the necessary integration of energy reflected at all angles is accomplished by the hemispherical surface. Only the determination of several minor correction factors remains, one accounting for the energy absorbed by the mirrored surface of the hemisphere, and a second

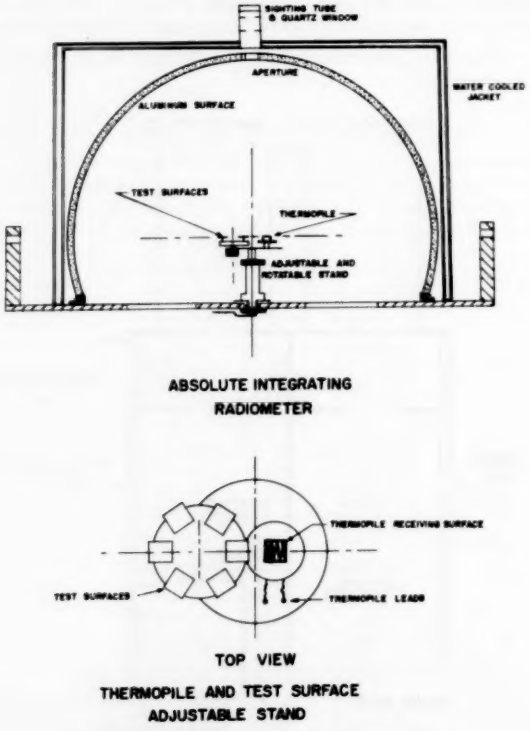


FIG. 2 ABSOLUTE INTEGRATING RADIOMETER

for the energy that is reflected off the test surface that passes out through the aperture. By taking the ratio of the energy reflected onto the thermopile from the test surface to that energy directly incident onto the thermopile and using the correction factors, one is able to determine the reflectivity.

Aluminum was chosen for the hemispherical surface coating because of its resistance to tarnishing and also for its relatively constant and high spectral reflectivity in the solar region of the spectrum. The hemisphere is covered with a cooling jacket to maintain defined conditions around the thermopile and test surfaces.

A fused quartz window, 2 mm in thickness, is installed over the orifice to eliminate any effects of convection currents on the thermopile reading. Throughout the important part of the solar spectrum, 0.3 to 3.0 microns, the fused quartz filters evenly and has a transmissivity of 92 per cent (6).

A Kipp and Zonen solarimeter, type G-19 thermopile, without base and screen was used in this instrument. The thermopile and test surfaces are mounted on a rotatable stand, Fig. 2. This stand can be rotated so that solar energy is directed either on the thermopile or test surface. The test-surface holder is provided with positions for six test surfaces. When measurements on one surface are completed, another surface can be rotated into the measuring position.

The complete instrument is mounted on a tripod so that the sun can be followed throughout a test day.

EXPERIMENTAL PROCEDURE

Comparison Method. The test apparatus is fixed on the sun so that the test surfaces are completely covered by solar energy. Both the reference and test surfaces are then allowed to come to their respective equilibrium temperatures. During this period, temperature readings are recorded as a function of time; continually adjusting the apparatus to keep the solar energy on the surfaces.

Two test procedures, designated as A and B, were used. One method, procedure A, consisted of raising the test-surface temperature up to that of the standard surface by electrical heating while both surfaces are exposed to the sun. In procedure

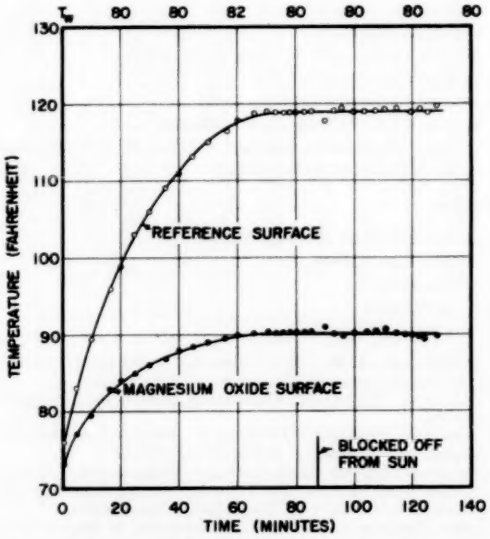


FIG. 3 TIME-TEMPERATURE EQUILIBRIUM CURVE-COMPARISON METHOD

B, both surfaces were allowed to come to their own equilibrium temperatures when exposed to solar radiation with no electrical heat supplied to either surface. Fig. 3 shows a typical time-temperature history using test procedure B.

After the surfaces have reached equilibrium using either procedure, the orifices at the top of the apparatus are blocked off from solar energy, but not from the atmosphere. In the case of procedure A, electrical energy is added in such amounts that the surfaces stay at the same equilibrium temperature as when exposed to solar energy. In procedure B, each surface is raised up to the equilibrium temperature it had when exposed to solar energy. A voltmeter and ammeter are used to measure these electrical inputs to the heating elements. The readings obtained are then corrected for power losses in the lead lines to the heaters.

Integrating-Radiometer Method. With the energy of the sun directly incident on the thermopile surface, the millivolt output is measured on a recording potentiometer. The test surface and thermopile are then rotated 180 deg such that the incident energy first falls on the test surface and is then reflected onto the thermopile surface; the output is again measured on the recording potentiometer. The final measurement is the determination of the amount of energy directly emitted by the test surface. This is accomplished by blocking off the orifice at the top of the instrument and recording the resulting millivolt output of the thermopile. The surface temperature of the test sample is also measured. Two correction factors are needed for determining the reflectivity: (*a*) the energy absorbed by the mirrored surface, and (*b*) the amount of energy that is reflected off the test surface and passes out through the aperture. The reflectivity and consequently the absorptivity of the hemispherical surface (evaporated aluminum film) is measured by putting a surface of the same coating as the hemisphere in the test position. The measurement of the reflectivity is then made as for any surface. An average of sixteen test runs was used to calculate this value and the reflectivity was found to be 0.90. The measurement of the energy loss through the aperture is determined in the following way: The incoming energy is allowed to fall upon the test surface and a millivolt reading is made. Now a surface of the same size as the aperture and coated with a mixture of acetylene soot and lampblack is placed next to the aperture. It is assumed that the energy absorbed by this surface is the same as the energy leaving through the aperture. A second reading on the recording potentiometer will now be smaller by the amount absorbed by the blackened surface. The energy lost through the aperture was determined by this procedure for each specimen and found to vary from 1 to 4 per cent.

Test Results and Discussion

The reliability of the present test apparatus was established by measuring the absorptivity of several surfaces for which previous values are available in the published literature. In the case of the comparative apparatus three such surfaces were checked and as shown in Table 1, the agreement is acceptable irrespective of whether test procedure A or B was used in the present investi-

Table 1 Total Absorptivity for Solar Radiation

		Experimental Results		
Materials	Surface Preparation	Previous Investigations	Comparative Methods*	Absolute Radiometer Method
Magnesium Oxide	Smooth, hard packed surface.	0.14, Ref. 3	0.14A 0.16B	
75ST Alclad Aluminum	Washed with Ivory soap for 3 minutes, wiped with benzol until no dirt showed on cloth.	0.59, Ref. 7	0.59A	
75ST Alclad Aluminum	Buffed on soft canvas wheel using tripoli from 2 to 3 minutes per 3 sq. inches. Hand rubbed for 4 to 5 minutes to high polish, washed with Ivory soap, wash with benzol.	0.34, Ref. 7	0.37A 0.36B	0.33
24SO Alclad Aluminum	Buffed on soft canvas. ..... (Same procedure as for 75ST Alclad).		0.34A 0.35B	0.35
Poroloy (18 - 8 stainless steel)	Washed with benzol and then with acetone.			
28% porosity			0.63B	
31% porosity			0.64B	0.63
43% porosity			0.66B	
47% porosity			0.66B-0.68B	0.64
Tyler (AISI Type 304 stainless steel)	Washed with benzol and then with acetone.			
28 x 500 SMR mesh				0.86
20 x 200 mesh				0.77
20 x 350 mesh				0.73

* "A" denotes Procedure A, while "B" denotes Procedure B as described in the text.

gation. For the integrating radiometer a single check was made using 75S-T Alclad aluminum and yielded a value in agreement with that reported by Wilkes (7) and, in addition, demonstrated that the results obtained with the radiometer are consistent with those obtained with the comparative device. Such consistency between the present two measurement techniques is further demonstrated (Table 1) in the case of the 24S-O Alclad aluminum surface and for the Poroloy materials.

The most significant measurements reported herein are for two types of porous materials presently being considered for transpiration cooling of high-speed vehicles or for gas-turbine blades which are exposed to hot gases. The first such material is designated by the trade name Poroloy (8) and is fabricated of stainless-steel wire wound on a mandrel, one layer over another, to yield the desired porosity, then finally sintered. To specify the surface further, photomicrographs were obtained as shown in Fig. 5, and a chemical analysis yielded the following proportions of chromium and nickel for the four Poroloy specimens tested:

Porosity of material	% Ni	% Cr
28 per cent	8.86	18.7
31 per cent	8.87	18.6
43 per cent	8.78	18.3
47 per cent	8.53	17.8

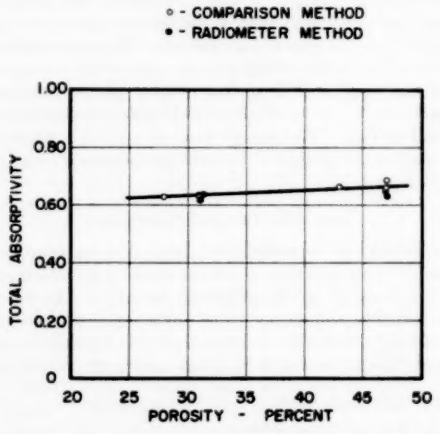


FIG. 4 TOTAL ABSORPTIVITY FOR SOLAR RADIATION OF POROLOY SURFACES

The resulting absorptivity values are shown in Table 1 and Fig. 4, where it is seen that they are quite high (0.63 to 0.68) and tend to increase with increasing porosity. Such high values are to be expected since the small openings on the surface act essentially as black bodies with a correspondingly high value of the absorptivity.

The second group of porous surfaces, designated as Tyler materials (9), are fabricated by weaving the wire to the desired mesh size, resulting in a rather dense screen-like surface.

Two of the surfaces are of similar weave but differ in one dimension of mesh size, while the third surface is designated SMR (sprayed, melted, and rolled) and is of a finer mesh size. Photomicrographs showing the construction of these materials are presented in Fig. 6. As a consequence of this type of construction there are many uniform depressions in the surface which result in increased absorptivity. The resulting values are between 0.73 and 0.86 and are given in Table 1.

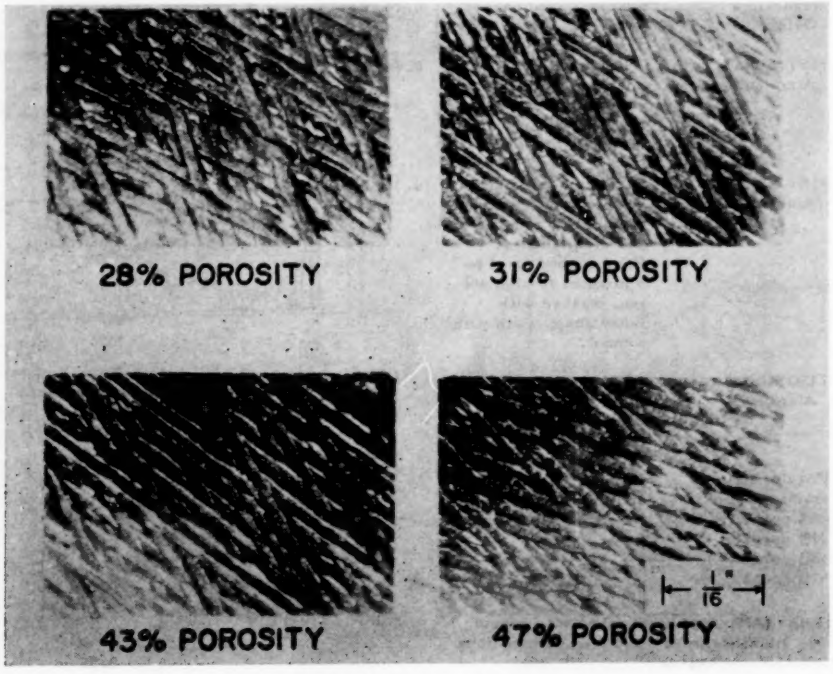


FIG. 5 POROLOY SURFACE MICROGRAPHS

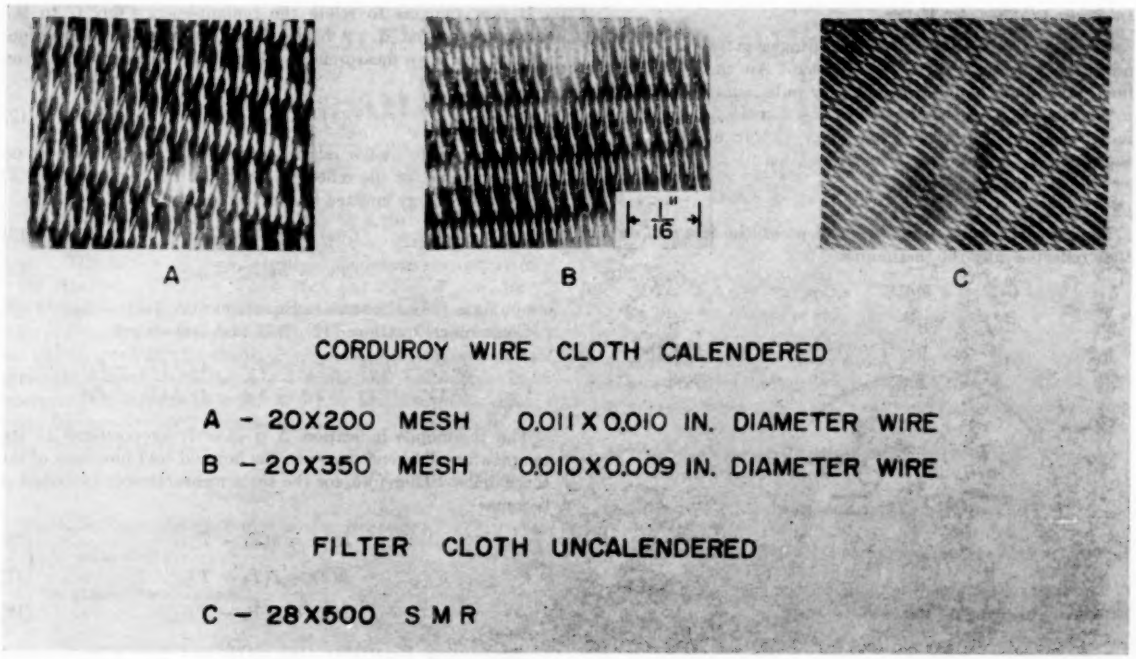


FIG. 6 TYLER SURFACE MICROGRAPHS

BIBLIOGRAPHY

1 "The Reflecting Powers of Rough Surfaces at Solar Wavelengths," by H. E. Beckett, Proceedings of the Physical Society of London, vol. 43, part 3, no. 238, May, 1931, pp. 227–241.

2 "The Experimental Determination of the Total Absorptivity of Several Important Engineering Materials for Solar Radiation," by Richard Birkebak, Master's thesis, University of Minnesota, Minneapolis, Minn., September, 1956.

3 "The Diffuse Reflecting Power of Various Substances," by W. W. Coblentz, National Bureau of Standards Bulletin, vol. 9, 1913, pp. 283–325.

4 "The Thermal Radiation Project" (final report), by R. V. Dunkle and J. T. Gier, University of California, Institute of Engineering Research, September, 1950, pp. 99–104.

5 "Measurement of Total Emissivity of Porous Materials in Use for Transpiration Cooling," by E. R. G. Eckert, J. P. Hartnett, and T. F. Irvine, Jr., *Jet Propulsion*, vol. 26, April, 1956, p. 280.

6 "Fused Quartz—Price Schedule," General Electric Catalog.

7 "Measurements of the Total Normal Emissivity of Materials," by G. B. Wilkes, Progress Report No. 9, Massachusetts Institute of Technology, Cambridge, Mass.

8 "Poroloy Catalog," Poroloy Equipment, Inc., Pacoema, Calif.

9 "Tyler Catalog," W. S. Tyler Company, Cleveland, Ohio.

Appendix

ANALYSIS OF DATA

A brief outline of the data analysis is given in this section. In the case of the comparison method, only test procedure B will be discussed as the extension to test procedure A is obvious.

Comparison Method—Test Procedure B

Reference-Surface Energy Balance. At steady state after being exposed to solar energy the following energy balance results

Energy in Energy out

$$\alpha_r J_i + \alpha_r' F_{rw} \sigma T_w^4 = \epsilon_r \sigma T_r^4 + q_{l_r} \quad \ldots\ldots [1]$$

Next the orifice at the top of the device is closed off to the incoming solar energy but not to the atmosphere. Electrical energy (q_{e_r}) is added in such an amount that the reference surface temperature is the same as when exposed to solar energy, all other conditions being equal

$$q_{e_r} + \alpha_r' \sigma T_w^4 = \epsilon_r \sigma T_r^4 + q_{l_r} \quad \ldots\ldots [2]$$

Combining Equations [1] and [2], we get for the incident solar energy J_i

$$J_i = \frac{q_{e_r} + \alpha_r' \sigma T_w^4 (1 - F_{rw})}{\alpha_r} \quad \ldots\ldots [3]$$

Test-Surface Energy Balance. The sample is allowed to come to its own equilibrium temperature when exposed to solar energy

Energy in Energy out

$$\alpha_s J_i + \alpha_s' F_{sw} \sigma T_w^4 = \epsilon_s \sigma T_s^4 + q_{l_s} \quad \ldots\ldots [4]$$

As before, the orifice at the top is closed off from solar energy and electrical energy q_{e_s} added in such amounts that the test surface assumes the same temperature as when exposed to solar energy

$$q_{e_s} + \alpha_s' \sigma T_w^4 = \epsilon_s \sigma T_s^4 + q_{l_s} \quad \ldots\ldots [5]$$

Combining Equations [3], [4], and [5] the absorptivity of the test surface becomes

$$\alpha_s = \alpha_r \frac{q_{e_s} + \alpha_s' \sigma T_w^4 (1 - F_{sw})}{q_{e_r} + \alpha_r' \sigma T_w^4 (1 - F_{rw})} \quad \ldots\ldots [6]$$

In the present case F_{sw} and F_{rw} are approximately 0.99 and the foregoing equation can be simplified with little error, resulting in the following expression for the absorptivity

$$\alpha_s = \alpha_r \frac{q_{e_s}}{q_{e_r}} \quad \ldots\ldots [7]$$

Integrating Radiometer Method

Three measurements are made during a given run to determine the reflectivity of a test surface. An energy balance is made on the thermopile for impinging radiation for each of these measurements. It is assumed that the thermopile receiving surface has an absorptivity of one (Fig. 7). Incident energy on thermopile when exposed to solar radiation

$$G(1) = J_i + \epsilon_t C^2 \rho_m{}^2 \rho_s{}' + e_s \rho_m C \ldots\ldots\ldots\ldots [8]$$

Next, for the radiant energy incident on the test surface and then reflected onto the thermopile

$$G(2) = J_i \rho_s \rho_m C + e_t C^2 \rho_m{}^2 \rho_s{}' + e_s \rho_m C \ldots\ldots\ldots [9]$$

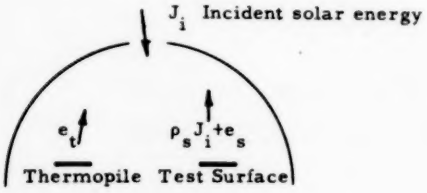


FIG. 7

The final measurement is made for the energy emitted by the sample by blocking off the sun

$$G(3) = e_t C^2 \rho_m{}^2 \rho_s{}' + e_s \rho_m C \ldots\ldots\ldots\ldots [10]$$

Solving Equations [8], [9], and [10] for ρ_s the reflectivity of the test surface

$$\rho_s = \frac{1}{\rho_m C} \frac{G(2) - G(3)}{G(1) - G(3)} \ldots\ldots\ldots\ldots [11]$$

It now remains to relate the radiant-energy flux G to the thermopile output Δ. A heat balance is written on the thermopile for the three measurements of energy. Incident energy on the thermopile

$$G(1) = \sigma \epsilon_t T_1{}^4 + q_l \ldots\ldots\ldots\ldots [12]$$

We may write similar relationships for the energy incident on the thermopile for the reflected energy off the test surface $G(2)$ and for the energy emitted by the test surface $G(3)$

$$G(2) = \sigma \epsilon_t T_2{}^4 + q_l \ldots\ldots\ldots\ldots [13]$$

$$G(3) = \sigma \epsilon_t T_3{}^4 + q_l \ldots\ldots\ldots\ldots [14]$$

assuming q_l to be the same in Equations [12], [13], and [14].

Combining Equations [12], [13], and [14] we get

$$\frac{G(2) - G(3)}{G(1) - G(3)} = \frac{T_2{}^4 - T_3{}^4}{T_1{}^4 - T_3{}^4} \approx \frac{4T_2{}^3 (T_2 - T_3)}{4T_3{}^3 (T_1 - T_3)} \ldots\ldots [15]$$

The thermopile indication Δ is directly proportional to the temperature difference between the hot and cold junctions of the thermopile. Therefore, for the three measurements indicated Δ becomes

$$\Delta(1) = k(T_1 - T) \ldots\ldots\ldots\ldots [16]$$

$$\Delta(2) = k(T_2 - T) \ldots\ldots\ldots\ldots [17]$$

$$\Delta(3) = k(T_3 - T) \ldots\ldots\ldots\ldots [18]$$

Substituting Equations [16], [17], and [18] into Equation [15], and combining with Equation [11], we arrive at the final relationship for the reflectivity of the test surface

$$\rho_s = \frac{1}{\rho_m C} \frac{\Delta(2) - \Delta(3)}{\Delta(1) - \Delta(3)} \ldots\ldots\ldots\ldots [19]$$

Similar Solutions for Free Convection From a Nonisothermal Vertical Plate[1]

By E. M. SPARROW[2] AND J. L. GREGG,[3] CLEVELAND, OHIO

An analysis is made for laminar free convection from a vertical flat plate having a nonuniform surface temperature. The following two families of surface temperatures were studied: (i) $T_w - T_\infty = Nx^n$, (ii) $T_w - T_\infty = Me^{mx}$. Both families submit to mathematical analysis by the conventional techniques of laminar boundary-layer theory; i.e., they permit the finding of similar solutions of the boundary-layer equations. Numerical solutions have been carried out for Pr = 0.7 and 1.0; i.e., for gases. Heat-transfer results are presented, as are temperature and velocity distributions.

Nomenclature

The following nomenclature is used in the paper:

b = plate width

C_1 = dimensional constant, $\left[\frac{g\beta N}{4\nu^2}\right]^{1/4}$

C_2 = dimensional constant, $\left[\frac{g\beta m}{4\nu^2}\right]^{1/4}$

c_p = specific heat at constant pressure

F = dimensionless dependent variable defined by Equation [7b]

G = dimensionless dependent variable defined by Equation [11b]

g = acceleration due to gravity

Gr_x = local Grashof number based on x, $\frac{g\beta(T_w - T_\infty)x^3}{\nu^2}$, dimensionless

Gr_L = over-all Grashof number based on L, $\frac{g\beta(\Delta T)L^3}{\nu^2}$, dimensionless, where ΔT is some arbitrarily selected temperature difference

h = local heat-transfer coefficient, $q/(T_w - T_\infty)$

$\bar{h}$ = average heat-transfer coefficient, $Q/Lb(\Delta T)$, where ΔT is some arbitrarily selected temperature difference

k = thermal conductivity

L = plate length

M = dimensional constant in Equation [2]

m = exponent in Equation [2]

N = dimensional constant in Equation [1]

n = exponent in Equation [1]

Nu_x = local Nusselt number, hx/k, dimensionless

Nu_L = over-all Nusselt number, $\bar{h}L/k$, dimensionless

Pr = Prandtl number, $\frac{\nu}{\alpha} = \frac{c_p\mu}{k}$, dimensionless

Q = over-all heat-transfer rate, $\int_0^L qb\,dx$

q = local heat-transfer rate per unit area

T = static temperature

u = velocity component in x-direction

v = velocity component in y-direction

x = co-ordinate measuring distance along plate from leading edge

y = co-ordinate measuring distance normal to plate

α = thermal diffusivity, $k/\rho c_p$

β = coefficient of thermal expansion, $-\frac{1}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_p$

η = dimensionless similarity variable defined by Equation [7a]

θ = dimensionless temperature, $T - T_\infty/T_w - T_\infty$

μ = absolute viscosity

ν = kinematic viscosity

ξ = dimensionless similarity variable defined by Equation [11a]

ρ = fluid density

φ = dimensionless temperature, $T - T_\infty/T_w - T_\infty$

ψ = stream function

Subscripts

w = wall conditions

∞ = ambient conditions

Introduction

In a great many technical applications, the surface from which heat is being transferred is nonisothermal. For forced convection, heat transfer from nonisothermal walls has been treated with success by a number of investigators; references (1) through (4)[4] are examples of an extensive literature. For free convection, there has been only a limited amount of analytical work. In particular, for free convection on a vertical flat plate, on which attention will be focused here, only the following analyses have appeared: Sparrow and Gregg (5) give results for the case of uniform surface heat flux based on numerical solutions of the differential equations of the boundary layer. Approximate solutions using integral methods were obtained by Sparrow (6) for the following particular variations of surface heat-flux and surface temperature

$$q = q_{x=0}[1 \pm A(x/L)^p]$$

and

$$(T_w - T_\infty) = (T_w - T_\infty)_{x=0}[1 \pm B(x/L)^p]$$

where A, B, and p are positive constants. Siegel (7) also used a similar integral method for the uniform heat-flux case.

In considering the nonisothermal wall problem, it is natural to examine carefully those families of wall-temperature variations which submit to mathematical analysis by the conventional techniques of laminar boundary-layer theory. Stated more pre-

[1] The material presented here is taken in part from Chapter 6 of a PhD thesis submitted to Harvard University by E. M. Sparrow (see Bibliography, reference 9).

[2] Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics. Assoc. Mem. ASME.

[3] Lewis Flight Propulsion Laboratory, National Advisory Committee for Aeronautics.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 23, 1956. Paper No. 57—SA-3.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper

cisely, the wall-temperature variations to be examined are those which give rise to similar solutions of the laminar boundary-layer equations. Two such families of wall-temperature variations to be studied here are

$$T_w - T_\infty = Nx^n \quad [1]$$

$$T_w - T_\infty = Me^{mx} \quad [2]$$

where N, M, n, m, and T_∞ are constants. For each of these families, numerical solutions of the laminar boundary-layer differential equations were carried out for Prandtl numbers of 0.7 and 1.0; i.e., for gases. Heat-transfer results based on these solutions are presented, as are temperature and velocity profiles. Those interested in results are invited to pass over the section on analysis.

In addition to the laminar-flow assumption, the state is taken to be steady and fluid-property variations are neglected, except for those density variations necessary to give a buoyancy force. The boundary-layer form of the conservation equations is assumed to apply.[5]

Analysis

Physical Model and Co-Ordinates. The physical model and the co-ordinate system are shown in an elevation view in Fig. 1. Two physical situations are shown which come within the scope of the analysis. The left-hand sketch depicts the case where the local wall temperature T_w, which may vary with x, exceeds the ambient temperature T_∞. Under these circumstances the free-convection motion is upward along the plate as shown.[6] The right-hand sketch shows the situation where T_w, which again may be a function of x, is lower than the ambient temperature T_∞. In this case, the fluid flow is downward along the plate as shown.

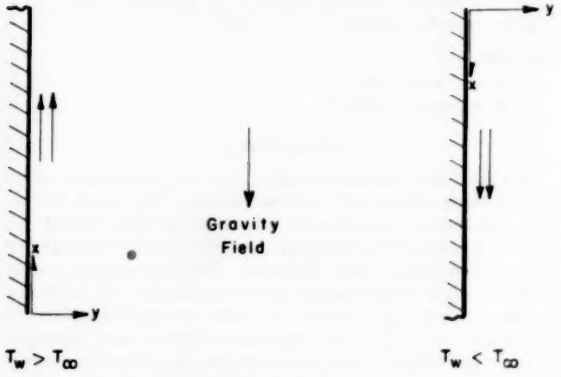


Fig. 1 Co-Ordinate Systems

If the co-ordinate systems are taken as indicated, the mathematical distinction between the two situations depicted in Fig. 1 vanishes when the conservation equations, as written later, are made dimensionless. So, separate analyses need not be made. Since it seems easier to visualize occurrences associated with the hot-wall case, i.e., $T_w > T_\infty$, the analysis and discussion will be directed toward this situation. However, the results will be presented in a manner applicable for both $T_w > T_\infty$ and $T_w < T_\infty$.

Conservation Laws. The equations expressing conservation of mass, momentum, and energy for steady laminar flow in a boundary layer on a vertical flat plate are, respectively, as follows

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \quad [3]$$

$$u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} = g\beta(T - T_\infty) + \nu\frac{\partial^2 u}{\partial y^2} \quad [4]$$

$$u\frac{\partial T}{\partial x} + v\frac{\partial T}{\partial y} = \alpha\frac{\partial^2 T}{\partial y^2} \quad [5]$$

In accord with the usual practice in free convection, the density has been considered a variable only in forming the buoyancy force $g\beta(T - T_\infty)$. Variations of all other properties are neglected. Viscous dissipation and work against the gravity field also have been neglected.

The boundary conditions appropriate to the problem are

$$\left.\begin{matrix} v = 0 \\ u = 0 \\ T_w = T_w(x) \end{matrix}\right\} y = 0 \qquad \left.\begin{matrix} u = 0 \\ T = T_\infty \end{matrix}\right\} y = \infty \quad [6]$$

The solution of Equation [3] as usual may be written in terms of a stream function ψ defined by the relations

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial x} \quad [3a]$$

Then the velocity components u and v in Equations [4] and [5] are replaced in favor of the stream function. From the substitution, there results the following pair of partial differential equations for ψ and T as functions of x and y

$$\frac{\partial \psi}{\partial y}\frac{\partial^2 \psi}{\partial x\,\partial y} - \frac{\partial \psi}{\partial x}\frac{\partial^2 \psi}{\partial y^2} = g\beta(T - T_\infty) + \nu\frac{\partial^3 \psi}{\partial y^3} \quad [4a]$$

$$\frac{\partial \psi}{\partial y}\frac{\partial T}{\partial x} - \frac{\partial \psi}{\partial x}\frac{\partial T}{\partial y} = \alpha\frac{\partial^2 T}{\partial y^2} \quad [5a]$$

Rather than deal directly with these two formidable partial differential equations, experience leads us to seek a way of transforming them to a pair of ordinary differential equations, which are easier to solve. In the usual terminology of boundary-layer theory, such a transformation is called a similarity transformation. It is not possible to carry out a similarity transformation of Equations [4a] and [5a] for any arbitrary wall-temperature variation. Two families of wall-temperature variations which do permit reduction of Equations [4a] and [5a] to ordinary differential equations have been given in Equations [1] and [2]. The similarity transformations and resulting ordinary differential equations for these wall-temperature variations are given in the following:

Transformation for $T_w - T_\infty = Nx^n$. A new independent variable η, called a similarity variable, is defined by

where

$$\left.\begin{matrix} \eta = C_1 y x^{\frac{n-1}{4}} \\ C_1 = \left[\dfrac{g\beta N}{4\nu^2}\right]^{1/4} \end{matrix}\right\} \quad [7a]$$

New dependent variables F and θ are given by

$$F(\eta) = \frac{\psi}{x^{\frac{n+3}{4}}}\frac{1}{4\nu C_1}, \qquad \theta(\eta) = \frac{T - T_\infty}{T_w - T_\infty} \quad [7b]$$

[5] Although the boundary-layer assumptions appear more difficult to justify for the nonisothermal-wall problem than for the isothermal-wall problem, their retention is a practical necessity.

[6] This statement applies to fluids showing the usual trend of density decreasing with increasing temperature.

The function θ is a dimensionless temperature and F is related to the velocities in the following way

$$\left.\begin{aligned} u &= 4\nu C_1^2 x^{\frac{n+1}{2}} F' \\ v &= -\nu C_1 x^{\frac{n-1}{4}} [(n+3)F + (n-1)\eta F'] \end{aligned}\right\} \quad \ldots\ldots [8]$$

The primes represent differentiation with respect to η.

Under the transformation Equations [7a] and [7b], the partial differential Equations [4a] and [5a] become

$$\left.\begin{aligned} F''' + (n+3)F''F - (2n+2)(F')^2 + \theta &= 0 \\ \theta'' + \Pr[(n+3)F\theta' - 4nF'\theta] &= 0 \end{aligned}\right\} \quad \ldots\ldots [9]$$

These are simultaneous ordinary differential equations which contain as parameters the Prandtl number and n (which specifies the shape of the wall-temperature variation). The coupling of the equations arises because the free-convection motion is due completely to temperature differences. The boundary conditions, Equations [6], transform to

$$\left.\begin{aligned} F &= 0 \\ F' &= 0 \\ \theta &= 1 \end{aligned}\right\} \eta = 0 \qquad \left.\begin{aligned} F' &= 0 \\ \theta &= 0 \end{aligned}\right\} \eta = \infty \quad \ldots\ldots [10]$$

Numerical solutions[7] of Equations [9] subject to the Boundary Conditions [10] have been carried out for Prandtl numbers of 0.7 and 1.0 (i.e., gases), for the following values of n: 3, 2, 1, 0.5, 0.2, 0, -0.2, -0.5, -0.8. Results of engineering interest based on these solutions will be given in a later section.

It is worth while noting that, because of the nature of the transformation given by Equation [7a], boundary conditions along the line $x = 0$ cannot be specified once the conditions at $y = 0$ and $y = \infty$ have been specified.[8] In fact, one must accept whatever conditions happen to be satisfied along $x = 0$ by the solution of Equations [9] subject to the boundary conditions [10]. For example, for $n > -1$ (which is true for all the numerical calculations made here), the velocity u is zero along $x = 0$.

Transformation for $T_w - T_\infty = Me^{mx}$, $m > 0$. First, consider the situation where the exponent m is positive. Then, new independent and dependent variables are introduced as follows by Equations [11a] and [11b], respectively

$$\left.\begin{aligned} \xi &= C_2 y (Me^{mx})^{1/4} \\ C_2 &= \left[\frac{g\beta m}{4\nu^2}\right]^{1/4} \end{aligned}\right\} \quad \text{where} \quad \ldots\ldots [11a]$$

$$G(\xi) = \frac{\psi}{(Me^{mx})^{1/4}} \frac{1}{4\nu C_2 m}, \qquad \varphi(\xi) = \frac{T - T_\infty}{T_w - T_\infty} \quad \ldots [11b]$$

ξ is usually termed a similarity variable, φ is a dimensionless temperature, and G is related to the velocities of the problem by the equations

$$u = \frac{4\nu C_2^2}{m} (Me^{mx})^{1/2} G', \qquad v = -\nu C_2 (Me^{mx})^{1/4} [G'\xi + G] \ldots [12]$$

The primes represent differentiation with respect to ξ. The transformation evidently does not apply for $m = 0$

Under the transformation Equations [11a] and [11b], the partial differential Equations [4a] and [5a] are reduced to the following ordinary differential equations

$$\left.\begin{aligned} G''' + GG'' - 2(G')^2 + \varphi &= 0 \\ \varphi'' + \Pr[G\varphi' - 4G'\varphi] &= 0 \end{aligned}\right\} \quad \ldots\ldots [13]$$

It is seen that the solution for G and φ depends upon the choice of Prandtl number. The absence of the exponent m from these equations is somewhat surprising.

The boundary conditions on G, φ, and their derivatives are identical to those for F, θ, and their derivatives given in Equation [10]. Numerical solutions for Equations [13] have been obtained for Prandtl numbers of 0.7 and 1.0.

As in the previous section, the nature of the transformation does not permit the specification of conditions along the line $x = 0$ once conditions at $y = 0$ and $y = \infty$ have been given. In this connection, it is worth while noting from Equations [12] and [11a] that the velocity u is not zero and not uniform (since G' varies with ξ) along $x = 0$.[9] It also may be seen that the temperature is not uniform along $x = 0$. These temperature and velocity conditions are certainly different from the uniform conditions along $x = 0$ usually encountered in boundary-layer analyses.

Transformation for $T_w - T_\infty = Me^{mx}$, $m < 0$. For negative values of m, the transformation used in the previous section fails because C_2, given in Equation [11a], is imaginary. One may find, without difficulty, a real transformation which will reduce Equations [4a] and [5a] to ordinary differential equations.[10] Study of these ordinary differential equations shows that their solutions have rather unusual characteristics. In particular, negative velocities, and/or temperatures less than ambient would be encountered. The cause of these unexpected findings may be traced directly to corresponding conditions along the line $x = 0$ which, because of the nature of the transformation, are not at our disposal.

It was deemed not worth while to carry out numerical integrations for this case because the results would have little practical value.

Results for $T_w - T_\infty = Nx^n$

The heat-transfer results will be presented first. Then, temperature and velocity profiles will be shown.

Local Heat Transfer. The local heat transfer from the surface to the fluid may be calculated using Fourier's law

$$q = -k\left[\frac{\partial T}{\partial y}\right]_{y=0}$$

Introducing the dimensionless variables from Equations [7a] and [7b], the expression for q becomes

$$q = -kNC_1 x^{\frac{5n-1}{4}} \left[\frac{d\theta}{d\eta}\right]_{\eta=0} \quad \ldots\ldots [14]$$

The derivative $[d\theta/d\eta]_{\eta=0}$, normally abbreviated as $\theta'(0)$, is a function of Prandtl number and of n, and is found from solutions of Equations [9].

The dependence of the local heat flux upon x is clearly seen from Equation [14]. Corresponding to a variation of $T_w - T_\infty$ proportional to x^n, there is a heat-flux variation proportional to $x^{\frac{5n-1}{4}}$. This information can be rephrased in another way. Suppose that a heat-flux variation proportional to x^r were prescribed, then the corresponding variation of $T_w - T_\infty$ would be proportional to $x^{\frac{4r+1}{5}}$. For the important case of uniform heat flux ($r = 0$), $T_w - T_\infty$ varies as $x^{1/5}$.

[7] The numerical integrations were carried out on an IBM Card Programmed Calculator using a technique presented in detail in appendix B of reference (8).

[8] This circumstance arises generally whenever a similarity transformation is used, no matter whether forced convection or free convection is being studied.

[9] The fact that u is nonzero at $x = 0$ may lead one to interpret solutions of Equations [13] as belonging to problems of combined, forced, and free convection.

[10] See Sparrow (9), chapter 6, for details.

Introducing the local heat-transfer coefficient, local Nusselt number, and local Grashof number as follows

$$h \equiv \frac{q}{(T_w - T_\infty)}, \quad \mathrm{Nu}_x \equiv \frac{hx}{k}, \quad \mathrm{Gr}_x \equiv \frac{g\beta|T_w - T_\infty|x^3}{\nu^2} \ldots [15]$$

the dimensionless representation of the local heat flux becomes

$$\frac{\mathrm{Nu}_x}{\mathrm{Gr}_x^{1/4}} = \frac{[-\theta'(0)]}{\sqrt{2}} \ldots [16]$$

The use of the absolute magnitude of the temperature difference in the Grashof number removes the necessity of separate consideration of $T_w > T_\infty$ and $T_w < T_\infty$, provided that the coordinates of Fig. 1 are used.

A plot of the Nusselt-Grashof relation given by Equation [16] is shown in Fig. 2 over the range $-0.8 \leq n \leq 3.0$ for Pr = 0.7 and 1.0. For values of n less than -0.6, it is seen that $\mathrm{Nu}_x/\mathrm{Gr}_x^{1/4}$ is negative. Physically, this corresponds to a heat transfer from the fluid to the wall, even though $T_w > T_\infty$. More will be said about this situation after the temperature profiles are shown.

It is of practical interest to inquire as to how well the heat flux for the nonisothermal wall could be predicted by local application of the isothermal-wall results. Such a procedure conceivably might be resorted to in the absence of other information about the variable wall-temperature problem. A comparison of local heat transfer for the isothermal and nonisothermal walls is given in Table 1. Except for very gradual variations of the wall temperature, i.e., n very close to zero, the procedure of locally applying isothermal-wall results to a nonisothermal wall appears unsatisfactory. A similar conclusion applies for forced convection over a flat plate for temperature variations given by Equation [1].[11]

TABLE 1 COMPARISON OF NONISOTHERMAL AND ISOTHERMAL HEAT TRANSFER

	$\frac{q \text{ (variable } T_w)}{q \text{ (constant } T_w)}$	
n	Pr = 0.7	Pr = 1.0
3	1.94	1.92
2	1.75	1.74
1	1.49	1.49
0.5	1.30	1.30
0.2	1.14	1.14
0	1	1
−0.2	0.81	0.81
−0.5	0.31	0.32
−0.8	−1.51	−1.68

Over-All Heat Transfer. The over-all heat transfer Q is also a quantity of engineering interest. Once the local heat flux is known from Equation [14], the over-all heat transfer can be calculated from

$$Q = b\int_0^L q\,dx \ldots [17]$$

where b is the plate width. For values of n less than -0.6, q is nonintegrable.[12] For $n > -0.6$

$$Q = 4bkNC_1L^{\frac{5n+3}{4}}\frac{[-\theta'(0)]}{5n+3} \ldots [17a]$$

For the wall-temperature variations considered here, there is no temperature difference which is characteristic of the problem; so, the choice of a temperature difference in defining an average heat-transfer coefficient (and hence an average Nusselt number) is purely arbitrary. The optimum circumstance would be to find a simple rule for choosing the temperature difference such that the same dimensionless representation of the heat-transfer results would apply no matter what the particular shape of the wall-temperature variation.

[11] See table 1 of reference (1).

[12] For $n = -0.6$, $\theta'(0) = 0$, so $q = 0$ for all x and hence $Q = 0$.

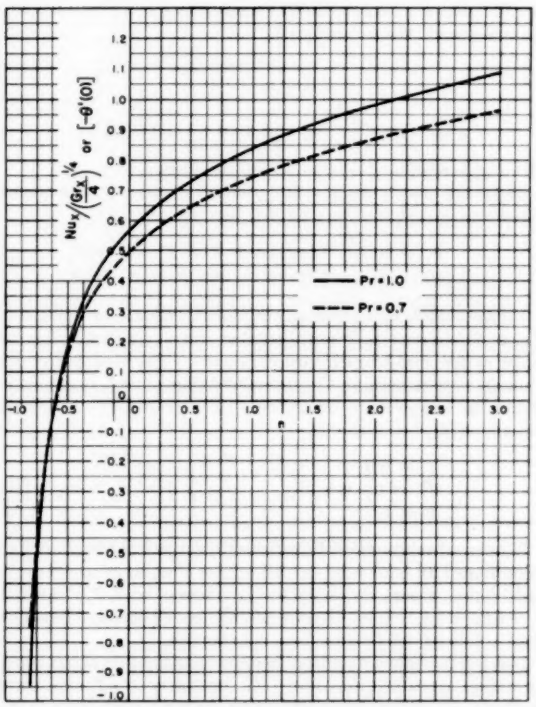


FIG. 2 PLOT OF $\mathrm{Nu}_x \Big/ \left(\frac{\mathrm{Gr}_x}{4}\right)^{1/4}$ AS A FUNCTION OF n FOR Pr = 0.7 AND Pr = 1.0. ($T_w - T_\infty = Nx^n$)

One might first consider using a mean temperature difference defined by

$$\overline{T_w - T_\infty} = \frac{1}{L}\int_0^L (T_w - T_\infty)dx \ldots [18]$$

The relationship between the average Nusselt number Nu_L and the over-all Grashof number Gr_L corresponding to such a temperature difference is shown in Table 2.

TABLE 2 $\mathrm{Nu}_L/\mathrm{Gr}_L^{1/4}$ BASED ON $\overline{(T_w - T_\infty)}$

	$\mathrm{Nu}_L/\mathrm{Gr}_L^{1/4}$	
n	Pr = 0.7	Pr = 1.0
3	0.856	0.962
2	0.757	0.849
1	0.629	0.707
0.5	0.555	0.626
0.2	0.506	0.573
0	0.471	0.534
−0.2	0.433	0.494
−0.5	0.370	0.428

It is seen that the Nusselt-Grashof relationship based on the temperature difference of Equation [18] depends strongly on the shape of the wall-temperature variation; i.e., on n.

A second choice, suggested by its simplicity, is to use the temperature difference halfway along the plate; i.e., $(T_w - T_\infty)_{L/2}$. The Nusselt-Grashof relationship corresponding to this choice of temperature difference is shown on Table 3, and a strong dependence on n may be observed.

No simple way of choosing the temperature difference has been found which will lead to a common relationship between Nu_L and Gr_L for all cases.

Temperature and Velocity Distributions. The dimensionless

TABLE 3 $Nu_L/Gr_L^{1/4}$ BASED ON $(T_w - T_\infty)L/3$

	$Nu_L/Gr_L^{1/4}$	
n	Pr = 0.7	Pr = 1.0
3	2.04	2.29
2	1.08	1.22
1	0.629	0.707
0.5	0.515	0.582
0.2	0.479	0.542
0	0.471	0.535
−0.2	0.481	0.549
−0.5	0.569	0.657

temperature $T - T_\infty / T_w - T_\infty$ is plotted against the similarity variable $\eta = \frac{y}{x}\left[\frac{Gr_x}{4}\right]^{1/4}$ in Fig. 3 for several representative values of n. Fig. 4 presents the dimensionless velocity

$$\frac{u}{[4g\beta|T_w - T_\infty|x]^{1/2}}$$

as a function of η for the same n values. Both figures are for Pr = 0.7.

From Fig. 3, it is seen that the temperature distributions for $n < 0$ differ notably in shape from that for $n = 0$ (the uniform wall temperature case). A readily visible inflection point occurs in the curve of $n = -0.5$, while the curve for $n = -0.8$ displays a "hill" where $T > T_w$.[13] For $n > 0$, the temperature distributions are similar in shape to that for $n = 0$. The shapes of the various velocity profiles in Fig. 4 do not exhibit gross differences such as those noted for the temperature profiles of Fig. 3.

For Pr = 1.0, one could present plots similar to those of Figs. 3 and 4. However, the dimensionless temperature and velocity profiles for Pr = 1.0 show the same dependence upon n as has al-

[13] Similar behavior has been shown by Schuh (2) for forced convection over a flat plate having a temperature variation given by Equation [1].

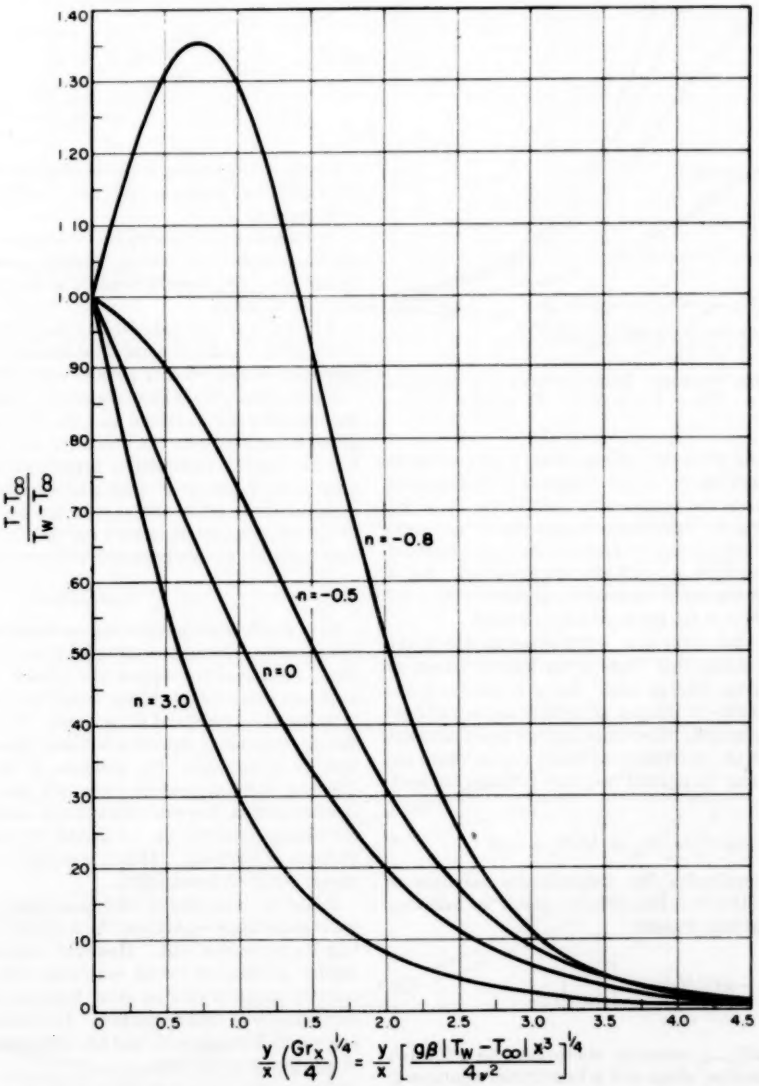


FIG. 3 DIMENSIONLESS TEMPERATURE DISTRIBUTIONS FOR SEVERAL VALUES OF n $(T_w - T_\infty = Nx^n)$. Pr = 0.7

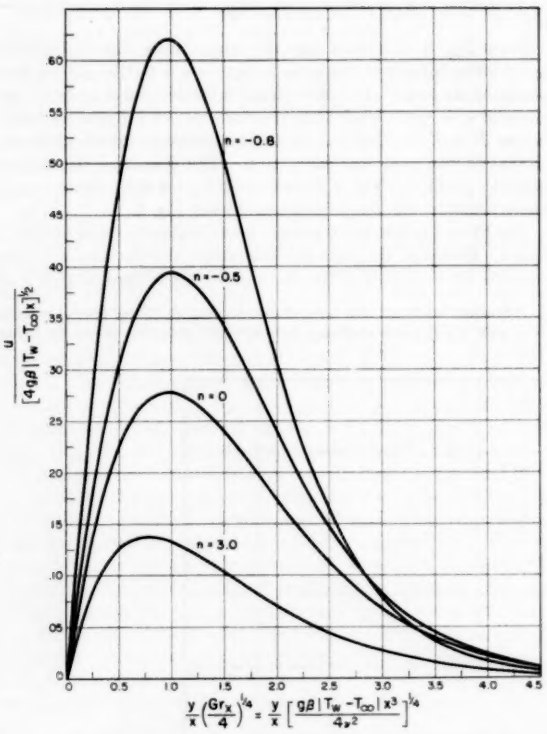


FIG. 4 DIMENSIONLESS VELOCITY DISTRIBUTIONS FOR SEVERAL VALUES OF n ($T_w - T_\infty = Nx^n$). Pr = 0.7

ready been displayed for Pr = 0.7. Hence, there is no need for so complete a presentation for Pr = 1.0. Typical curves showing the relationship between the temperature profiles for Pr = 0.7 and 1.0 are shown in Fig. 5. The relative orientation of the curves for $n = 0.5$ is typical for all other values of n studied, except for $n = -0.8$. So, curves for $n = -0.8$ also are presented. Fig. 6 shows the relative orientation of the velocity profiles for Pr = 0.7 and 1.0; the curves for $n = 0.5$ are typical for all cases.

Cases Where $n < -0.6$. For $n < -0.6$, investigation of the mathematical model shows that there is an infinite source of energy in the fluid at the leading edge. Being so endowed, the fluid (in the model) is able to transfer an infinite amount of heat to the plate in a finite length. Since such sources could not exist in nature, there arises an uncertainty as to the region where the results for $n < -0.6$ may be applied (i.e., how far from the leading edge).

RESULTS FOR $T_w - T_\infty = Me^{mx}$, $m > 0$

Heat Transfer. Introducing the dimensionless variables of Equations [11*a*] and [11*b*] into Fourier's law yields the following expression for the local heat transfer

$$q = -kC_2(Me^{mx})^{3/4}\left[\frac{d\varphi}{d\xi}\right]_{\xi=0} \quad \ldots\ldots\ldots\ldots [19]$$

The derivative $(d\varphi/d\xi)_{\xi=0}$, normally abbreviated as $\varphi'(0)$, is a function of Prandtl number alone and is found from solutions of Equations [13]. The information, given in Equation [19] allows a rephrasing of the problem in terms of a prescribed heat flux proportional to e^{mx} and a resulting wall-temperature variation proportional to $e^{3/4mx}$.

Using the definitions of Equation [15], the dimensionless form of the local heat flux is found to be

$$\frac{Nu_x}{Gr_x^{1/4}} = (mx)^{1/4}\frac{[-\varphi'(0)]}{\sqrt{2}},\ m > 0 \quad \ldots\ldots\ldots [20]$$

The calculated values of $[-\varphi'(0)]$ for Pr = 0.7 and 1.0 are 0.735 and 0.823, respectively.

The over-all heat transfer may be found by integrating Equation [19], using Equation [17]. Since there is no temperature difference characteristic of the problem, the definition of an average heat-transfer coefficient is completely arbitrary. Since there seems to be no advantage in using an average heat-transfer coefficient here, the results will simply be given in the following convenient dimensionless form

$$\frac{Q}{kb(T_w - T_\infty)_L}\left[\frac{g\beta(T_w - T_\infty)_L L^3}{\nu^2}\right]^{-1/4} = \frac{0.8}{\sqrt{2}}[-\varphi'(0)](mL)^{-3/4} \quad \ldots\ldots [21]$$

Uncertainties arising in the application of the heat-transfer results given by Equations [19], [20], and [21] will be discussed in a later section.

Temperature and Velocity Distributions. Dimensionless plots of the temperature and velocity profiles appear in Figs. 7 and 8, respectively. The Prandtl number is seen to effect no significant changes in shape.

When $x = 0$ is introduced into the ordinate and abscissa variables of Fig. 7 and 8, it is found (as already noted) that nonuniform temperature and velocity profiles exist at this location.

Limitations. One is immediately led to ask how strongly are the heat-transfer results bound up with the conditions which are imposed along $x = 0$ by the similarity transformation. This question can only be answered by experiment or by an analysis which permits the existence of other conditions along $x = 0$. Since such experiments or other analyses do not presently exist, the utility of the solutions obtained here for the exponential wall-temperature variation cannot be stated with certainty.

CONCLUSION

It is worth while mentioning an important difference between the variable wall-temperature problems in forced and free convection. For forced convection, the velocity and temperature problems are independent, either when the properties are constant or when $\rho\mu$ = const and ρk = const. Under either of these conditions, the energy equation is linear, provided that the Prandtl number is constant. So, solutions of the energy equation for different wall-temperature variations may be superposed.[14] In free convection, however, the velocity and temperature problems are always interrelated, no matter whether the properties are variable or constant. Hence, superposition of solutions of the energy equation is not valid.

It has not been proved that Equations [1] and [2] are the only wall-temperature variations which permit a similarity transformation to be carried out. However, experience with the rather similar problem of forced convection with variable free-stream velocity suggests that no other temperature variations will permit a similarity transformation. For wall-temperature variations other than Equations [1] and [2], different and less exact methods of solution must be used.

[14] This characteristic is the basis of the analyses of Chapman and Rubesin (3) and of Lighthill (4).

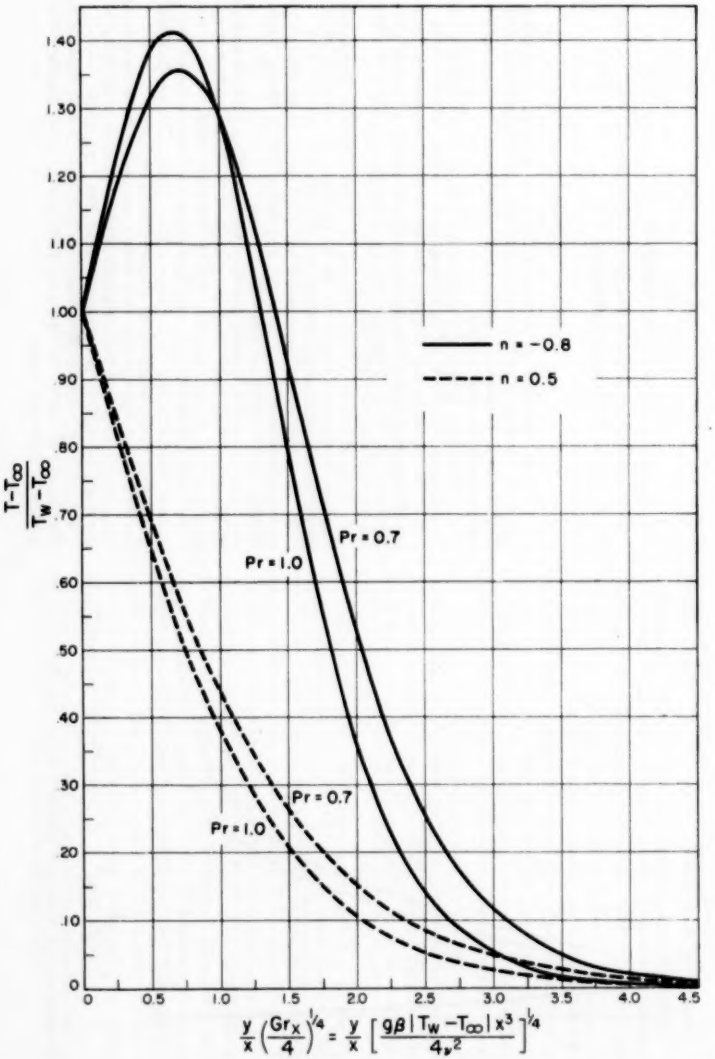


FIG. 5 TYPICAL CURVES SHOWING RELATIONSHIP BETWEEN DIMENSIONLESS TEMPERATURE DISTRIBUTIONS FOR Pr = 0.7 AND Pr = 1.0. RELATIVE ORIENTATION OF CURVES FOR $n = 0.5$ IS TYPICAL FOR OTHER CASES IN RANGE $-0.5 \leq n \leq 3.0$. $(T_w - T_\infty = Nx^n)$

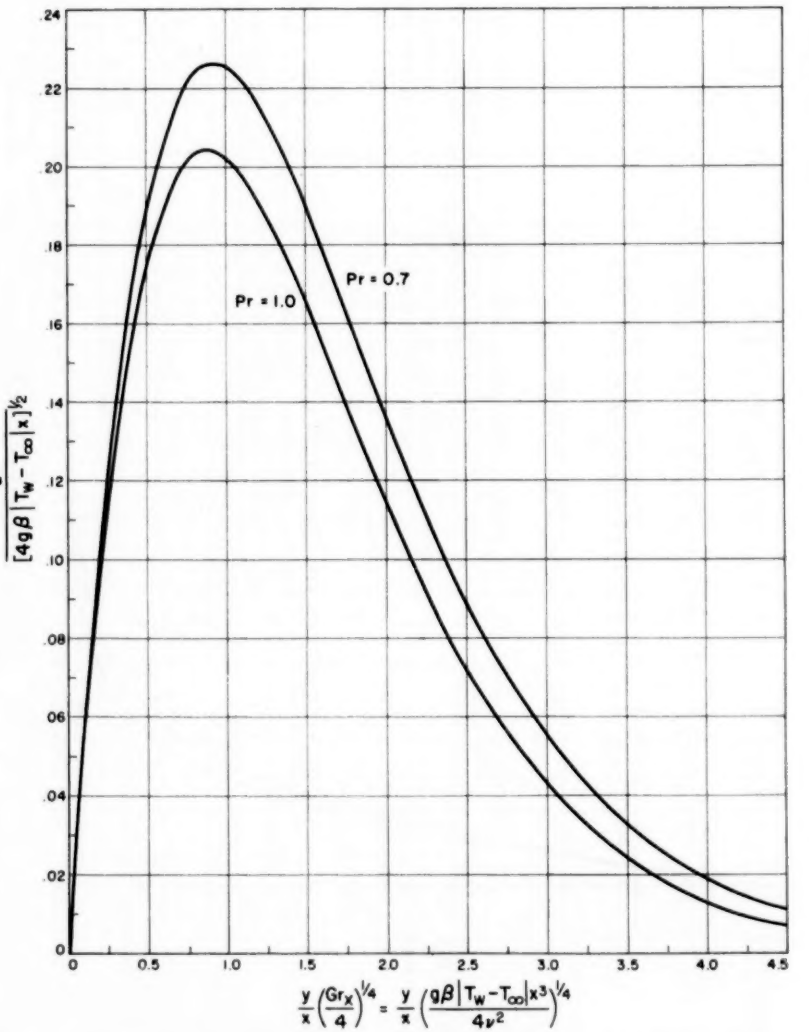


FIG. 6 TYPICAL CURVES SHOWING RELATIONSHIP BETWEEN DIMENSIONLESS VELOCITY DISTRIBUTIONS FOR Pr = 0.7 AND Pr = 1.0. RELATIVE ORIENTATION SHOWN FOR CURVES FOR $n = 0.5$ IS TYPICAL FOR OTHER CASES STUDIED. $(T_w - T_\infty = Nx^n)$

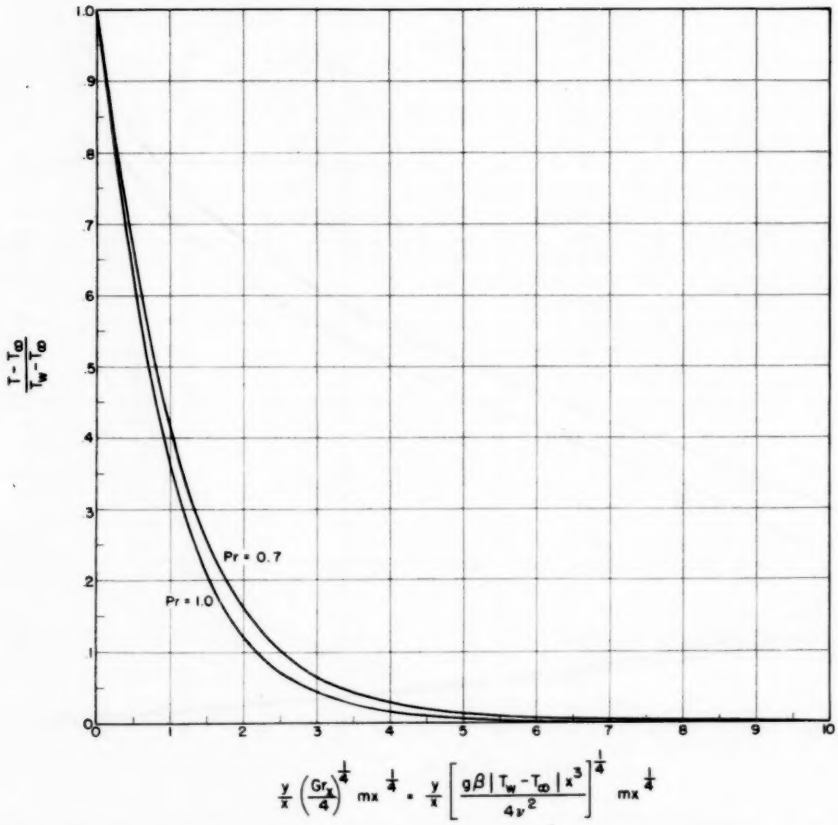


FIG. 7 DIMENSIONLESS TEMPERATURE DISTRIBUTIONS FOR $T_w - T_\infty = Me^{mx}$ FOR Pr = 0.7 AND 1.0

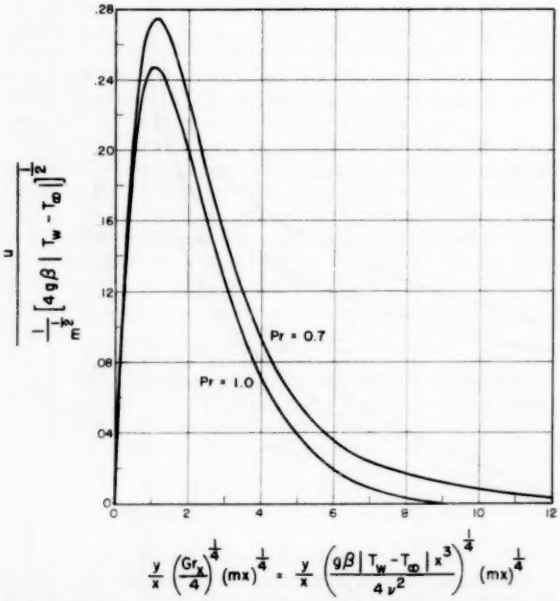


FIG. 8 DIMENSIONLESS VELOCITY DISTRIBUTIONS FOR $T_w - T_\infty = Me^{mx}$ FOR Pr = 0.7 AND 1.0

ACKNOWLEDGMENT

It is a pleasure to acknowledge the guidance of Prof. Howard W. Emmons of Harvard University.

BIBLIOGRAPHY

1 "Heat Transfer to Constant-Property Laminar Boundary-Layer Flows With Power-Function Free-Stream Velocity and Wall-Temperature Variation," by S. Levy, *Journal of the Aeronautical Sciences*, vol. 19, 1952, p. 341.

2 "Boundary Layers of Temperature," by H. Schuh, Reports and Translations No. 1007, AVA Monographs, British M.A.P., 1948.

3 "Temperature and Velocity Profiles in the Compressible Laminar Boundary Layer With Arbitrary Distribution of Surface Temperature," by D. R. Chapman and M. Rubesin, *Journal of the Aeronautical Sciences*, vol. 16, 1949, p. 547.

4 "Contributions to the Theory of Heat Transfer Through a Laminar Boundary Layer," by M. J. Lighthill, Proceedings of the Royal Society of London, series A, vol. 202, 1950, p. 359.

5 "Laminar Free Convection From a Vertical Plate With Uniform Surface Heat Flux," by E. M. Sparrow and J. L. Gregg, Trans. ASME, vol. 78, 1956, pp. 435–440.

6 "Laminar Free Convection on a Vertical Plate With Prescribed Nonuniform Wall Heat Flux or Prescribed Nonuniform Wall Temperature," by E. M. Sparrow, NACA TN 3508, 1955.

7 "Analysis of Laminar and Turbulent Free Convection From a Smooth Vertical Plate With Uniform Heat Dissipation per Unit Surface Area," by R. Siegel, G. E. Report R54GL89, 1954.

8 "An Analysis of Laminar Free-Convection Flow and Heat Transfer About a Flat Plate Parallel to the Direction of the Generating Body Force," by S. Ostrach, NACA TR 1111, 1953.

9 "Free Convection With Variable Properties and Variable Wall Temperature," by E. M. Sparrow, PhD thesis, 1956, Harvard University, Cambridge, Mass.

Laminar Mass and Heat Transfer From Ellipsoidal Surfaces of Fineness Ratio 4 in Axisymmetrical Flow[1]

By SHAO-YEN KO[2] and H. H. SOGIN[3]

The calibration of the heat-mass analog given by Sogin (7)[4] is employed to obtain mean coefficients of heat transfer from the nosepieces of ellipsoid-cylinders to air in axisymmetrical flow. The results for the ellipsoidal surface of axis ratio 4:1 are represented by the equation

$$\left(\frac{h}{Gc_p} N_{Pr,f}^{2/3}\right)\left(\frac{GS}{\mu_f}\right)^{1/2} = 0.76$$

in the range of the Reynolds number from 32,500 to 280,000. They are compared with the results of a wedge-flow approximation of the boundary-layer solution and with results from other investigations on related shapes.

Nomenclature

The following nomenclature is used in the paper:

A = area
a = semi-minor axis of the generating ellipse
B_1, B_2 = constants
b = mean coefficient of mass transfer
$C, \bar{C}$ = constant; mean value of constant from several tests
c_p = specific heat of air at constant pressure
D = diameter
$\mathfrak{D}$ = diffusivity of vapor in air
G = mass velocity of main stream
K_1 = constant defined in Equation [6]
h = coefficient of heat transfer by convection
k = thermal conductivity of air
L = major axis of generating ellipse
M = molecular weight
m = rate of mass transfer
$N_{Pr} = \mu_f c_p / k_f$
$N_{Re} = GS/\mu_f$
$N_{Sc} = \nu/\mathfrak{D}$
p = pressure
R = gas constant
r = radius
S = total length of semi-ellipsoidal surface measured from stagnation point along meridian profile = $1.0723L$ for $L/a = 4$
T = absolute temperature
U = stream velocity
x = distance measured from stagnation point along meridian profile
ϵ = jet-velocity correction as defined in Equation [12]
ζ = a coefficient defined in Equation [13]
η_r = recovery factor
λ = latent heat of sublimation
μ = dynamic viscosity
ν = kinematic viscosity
ρ = density or concentration

Subscripts

a = air
f = mean film
j = jet stream
s = local value at outer edge of boundary layer
v = vapor
x = local value at distance x
w = wall or surface
∞ = free stream

Introduction

Calculation of the heat transfer on blunt surfaces of revolution is encountered in the design of anti-icing equipment for some types of aircraft. When total rates of heat transfer are needed, the mean coefficient of heat transfer from an isothermal surface may provide a satisfactory estimate. The purpose of the present investigation has been to provide such data.

Mean coefficients of mass transfer on the nosepieces of ellipsoid-cylinders in axisymmetrical flow were measured in a 7-in-diam open-air jet. The axis ratio of the ellipsoids was 4:1, and the models, made of naphthalene, were of two sizes, their diameters being 1 and 2 in. The experimental procedure, the reduction of the data, and the transliteration to a heat-transfer correlation were essentially the same as those described by Sogin (7). The Reynolds number based on profile length ranged from 32,500 to 280,000, and the results indicated that the transfer was entirely laminar.

The results of a wedge-flow approximation to the boundary-layer solution are found to be 10 per cent less than the experimental values. The calculations are based on the assumption of incompressible flow and constant fluid properties. The heat transfer at the stagnation point of the ellipsoid-cylinder is calculated after Reshotko and Cohen (5). The rest of the surface is reduced to two dimensions by means of Mangler's transformation, and the local heat-transfer coefficients are determined with wedge-flow approximations after Eckert and Livingood (1). The coefficients are then transformed to the corresponding values in axisymmetry, and the results are integrated to obtain mean coefficients.

Finally, the results of the present investigation are compared with the experimental results of Stalder and Nielsen (8) on the

---

[1] This paper is part of a dissertation presented by the senior author for the degree of Doctor of Philosophy at Illinois Institute of Technology, Chicago, Ill., 1955.

[2] Senior Research Engineer, Minneapolis-Honeywell Regulator Company, Minneapolis, Minn. Assoc. Mem. ASME.

[3] Assistant Professor of Engineering, Brown University, Providence, R. I. Assoc. Mem. ASME.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 1, 1957. Paper No. 57—SA-44.

hemisphere-cylinder and of Lewis and Ruggeri (3) on the ellipsoid-cylinder of 3:1 fineness ratio.

Experimental Apparatus and Procedure

Because the experimental method has been described with considerable detail in (7), only the essential differences are noted here. The configuration and some nomenclature of the present investigation are shown in Fig. 1. The air stream issued from a 7-in-diam nozzle, and the outlet plane of the nozzle was about 10 in. from the model.

The molds were made of dental stone using brass patterns. The mold surfaces were sealed with a coating of Shell Epon 828 resin. The mold for the 1-in. specimens was a unified cavity, and the 2-in. mold was split. Brass skeletal structures, bored to distribute the molten naphthalene to all parts of the molds, were suitably equipped for mounting the specimens in the air streams.

On the average, the total mass transferred was about 100 and 400 milligrams for the 1 and 2-in. specimens, respectively, and the maximum error in the weight measurements was about 3 per cent of the total mass transferred.

During each run measurements were made to determine the temporal mean stagnation temperature T_0, and core, or center line velocity U_j. The temperatures ranged from 62 to 86 F, and the velocities from 27 to 130 fps.

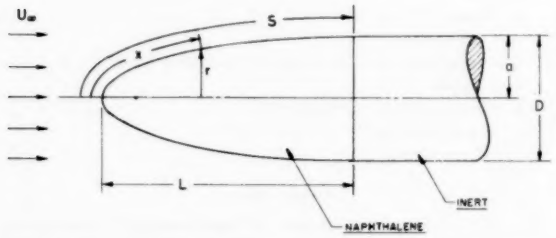


Fig. 1 Ellipsoidal Surface of Fineness Ratio 4 in Axisymmetrical Flow

The tests are divided into two series, Series A and B referring respectively to the 1 and 2-in. models.

Reduction of Data

Because the method of reducing the data is somewhat different from that in reference (7), it is presented here in detail.

For zero concentration of naphthalene in the main stream, a local coefficient of mass transfer b_x is defined by the equation

$$m_x'' = b_x \frac{p_{vw,x}}{R_v T_{w,x}} \quad \ldots\ldots [1]$$

The partial pressure $p_{vw,x}$ is the saturation pressure corresponding to the temperature $T_{w,x}$. They are related by the equation

$$\log_{10} p_{vw,x} = B_1 - \frac{B_2}{T_{w,x}} \quad \ldots\ldots [2]$$

where B_1 and B_2 are constants.

Since T_w is not measured directly, an additional relationship between p_{vw} and T_w is required. It is obtained from a heat balance as follows:

As is usually done, it is assumed that the conduction in the solid and the radiation from the surroundings are negligibly small compared with other modes of heat transfer. Then the rate of heat transfer by convection (including the effect of aerodynamic heating) to any place on the surface is equal to the rate of heat loss by sublimation. Thus

$$h_x \left(T_s + \eta_r \frac{U_s^2}{2c_p} - T_{w,x} \right) = b_x \frac{p_{vw,x}}{R_v T_{w,x}} \lambda \quad \ldots\ldots [3]$$

From the heat-mass analog, in the case that $\rho_v \ll \rho_a$

$$b_x = \frac{h_x}{\rho_a c_p} \left(\frac{N_{Pr}}{N_{Sc}} \right)^{2/3} \quad \ldots\ldots [4]$$

Placing $\eta_r = (N_{Pr})^{1/2}$ and approximating ρ_a with $p_0/R_a T_w$, we obtain

$$T_{w,x} = T_s + (N_{Pr})^{1/2} \frac{U_s^2}{2c_p} - K_1 \frac{p_{vw,x}}{p_0} \quad \ldots\ldots [5]$$

where

$$K_1 = \frac{M_v}{M_a} \left(\frac{N_{Pr}}{N_{Sc}} \right)^{2/3} \frac{\lambda}{c_p} \quad \ldots\ldots [6]$$

For naphthalene-air diffusion $N_{Sc} = 2.5$ and $K_1 = 1860$ F.

After introduction of the stagnation temperature, Equation [5] becomes

$$T_{w,x} = T_0 - \frac{U_j^2}{2c_p} [1 - (N_{Pr})^{1/2}] \left(\frac{U_s}{U_j} \right)^2 - K_1 \frac{p_{vw,x}}{p_0} \quad \ldots [7]$$

This equation involves local temperature, pressure, and velocity, but mean values are needed because total rates of mass transfer are measured. Hence, after multiplying Equation [7] through by $2\pi r\, dx$, integrating from $x = 0$ to $x = S$, and dividing through by the area A, we obtain the mean wall temperature

$$T_w = T_0 - \frac{U_j}{2c_p} [1 - (N_{Pr})^{1/2}] \frac{\int_0^S \left(\frac{r}{L} \right) \left(\frac{U_s}{U_j} \right)^2 d\left(\frac{x}{L} \right)}{0.2014} - K_1 \frac{p_{vw}}{p_0} \quad \ldots [8]$$

The constant 0.2014 is the ratio of the ellipsoidal area to $2\pi L^2$. The ratio U_s/U_j has been taken from references (4) and (6), and the value of the integral is 1.104. Hence

$$T_w = T_0 - 10^{-3} \times 0.0144 U_j^2 - 0.9 p_{vw} \quad \ldots [9][8]$$
$$\text{(I)} \quad - \quad \text{(II)} \quad - \quad \text{(III)}$$

where U = fps, p = psf, and T = deg R. It may be noted that each of the two corrective terms, one for the net effect of the jet expansion and the aerodynamic heating and the other for the cooling due to sublimation, are of the order 0.1 deg F. In a first approximation they could be neglected, but in a second approximation they should be taken into account, particularly at the higher speeds and higher temperatures.

Equations [2] and [9] were solved simultaneously for p_{vw} and T_w in each run. Thus it was tacitly assumed that since the variation of T_w was small Equation [2] could be used to relate the mean values.

Mean values of the mass-transfer coefficient, denoted with b, were then calculated with Equation [1] in the form

$$b = \frac{R_v T_w}{p_{vw}} \frac{m}{A} \quad \ldots\ldots [10]$$

Finally, the product

$$\left(\frac{b}{U_j} N_{Sc}^{2/3} \right) \left(\frac{U_j S}{\nu} \right)^{1/2} \equiv C \quad \ldots\ldots [11]$$

[8] The Roman numerals are used for reference to the respective terms.

TABLE 1 SUMMARY OF RESULTS OF THE INDIVIDUAL SERIES

Series	Calculation*	$\bar{C}$	$\frac{\text{Max}\lvert C-\bar{C}\rvert}{\bar{C}}$	Standard Deviation of C
A	(i)	0.749	--	0.0196
(1 in. diam.,	(ii)	0.788	0.053	0.0167
19 runs)	(iii)	0.800	0.048	0.0057
B	(i)	0.712	--	0.0408
(2 in. diam.,	(ii)	0.772	0.075	0.0292
22 runs)	(iii)	0.785	0.061	0.0293

*(i) (I) in Eq. 9 and values of B_1, B_2 in Eq. 2 taken from [2].

(ii) (I) and $B_1 = 11.884$, $B_2 = 6713$ from [9].

(iii) (I)-(II)-(III) and B_1, B_2 as in (ii).

was calculated for each run. For laminar transfer on a surface of given geometry this product is presumed to be a constant. In fact, although different mean values (denoted by $\bar{C}$) were found for the two series, the deviations from the mean in each case were small enough to retain Equation [11] as the final form of the correlating equation. The results are shown in Table 1.

Discussion of Results

The values of $\bar{C}$ have been calculated in three ways, which are described below Table 1. Comparing the first two results in each series and assuming that the reduction of the standard deviations are significant, we conclude that the values of B_1 and B_2 adopted in reference (7) from (9) are more accurate than those found in reference (2). Further evidence substantiating this choice is that the values of $\bar{C}_A$ and $\bar{C}_B$ have been brought into better agreement.

Introducing the correction terms (II) and (III) of Equation [9] reduces the scattering significantly in Series A and hardly affects it in Series B, possibly as a result of the better reproducibility achieved in preparing and handling the smaller specimens. Thus there is some indication that the corrections for aerodynamic heating and for cooling by sublimation are significant and should be taken into account; this has not been done by previous investigators who have performed this type of experimentation.

The final difference between $\bar{C}_A$ and $\bar{C}_B$ is ascribed to a scale effect, the fact that the ratios of the jet diameter to the specimen diameter in the two series are unequal. This difference is now eliminated by introducing a solid-blockage correction, which transforms the jet velocity U_j to the free-stream velocity U_∞. The correction for an open jet is of the form

$$U_\infty = U_j(1 - \epsilon) \quad [12]$$

where

$$\epsilon = \zeta \left(\frac{2a}{D_j}\right)^2 \quad [13]$$

The coefficient ζ is a constant, the same in both Series A and B. With the foregoing assumption it may be shown readily that

$$\zeta = \frac{\bar{C}_B^2 - \bar{C}_A^2}{\bar{C}_B^2\left(\frac{2a_B}{D_j}\right)^2 - \bar{C}_A^2\left(\frac{2a_A}{D_j}\right)^2} \quad [14]$$

Hence, $\zeta = 1.81$, $\epsilon_A = 0.0528$, and $\epsilon_B = 0.00422$. It follows that for both series, Equation [11] becomes

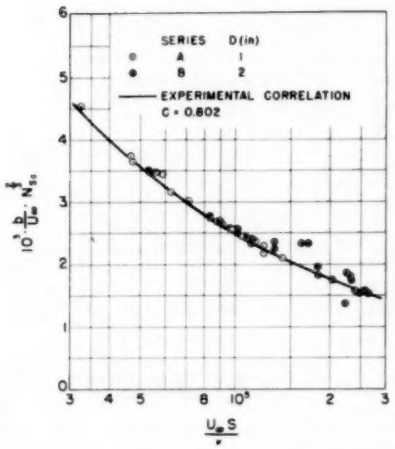


FIG. 2 EXPERIMENTAL RESULTS ON SUBLIMATION FROM ELLIPSOIDAL SURFACES

$$\left(\frac{b}{U_\infty} N_{Sc}^{2/3}\right)\left(\frac{U_\infty S}{\nu}\right)^{1/2} = 0.802 \quad [15]$$

The curve of this equation is shown on the semi-logarithmic plot of Fig. 2 together with the experimental points.

Transliterating Equation [15] to the corresponding case of heat transfer, as was done in reference (7), we obtain

$$\left(\frac{h}{Gc_p} N_{Pr,f}^{2/3}\right)\left(\frac{GS}{\mu_f}\right)^{1/2} = 0.802 \quad [16]$$

Subscript f refers to the mean film temperature, and G is the mass velocity of the free stream.

Theoretical Calculation

The method of the calculation is described and the results are presented; the details are omitted to conserve space.

The Prandtl number was set equal to 0.7, and the surface temperature was assumed uniform. The difference between the surface and fluid temperatures was taken to be small and the flow incompressible so that constant values of the fluid properties could be employed.

The velocity roof over the ellipsoid-cylinder was based on the calculations of McNown and Hsu (4) and on the experiments of Rouse and McNown (6), except that the velocities at the region of the stagnation point were interpolated with those of an ovary ellipsoid of 4:1 axis ratio in irrotational flow. The heat transfer at the stagnation point was calculated after Reshotko and Cohen (5). The remainder of the surface was reduced to two dimensions by means of Mangler's transformation, and the local coefficients were determined with simple wedge-flow approximations after Eckert and Livingood (1). The coefficients were then transformed to the corresponding values in axisymmetry. These results are shown in Fig. 3.

Finally, the local values were integrated over the area of the ellipsoidal nosepiece with the result that Equation [16] would apply if the constant were 0.721. This value is 10 per cent less than the experimental value. According to Eckert and Livingood, who compared a number of solutions, a difference of this order may be expected on account of the approximative nature of the calculation. Further, since the experimental value may be somewhat high on account of the turbulence in the free jet the agreement is considered satisfactory, and a compromise at the value 0.76 is suggested for applications.

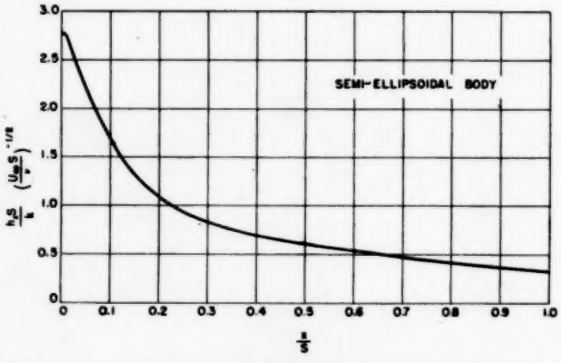


FIG. 3 LOCAL COEFFICIENTS OF HEAT TRANSFER ON ELLIPSOID-CYLINDER OF FINENESS RATIO 4

COMPARISON WITH OTHER INVESTIGATIONS

The experimental result of the present investigation is compared with the mean coefficients of heat transfer on some related surfaces. For this purpose it is convenient to write Equation [16] in the form

$$\left(\frac{h}{Gc_p} N_{Pr,f}^{2/3}\right)\left(\frac{GD}{\mu_f}\right)^{1/2} = f(L/a) \quad \ldots\ldots\ldots [17]$$

At the present time there are data for only a few fineness ratios and the complete shape of $f(L/a)$ cannot be given now, but the few known values are presented in Table 2.[6] It is understood that the presentation is limited to completely laminar transfer on blunt surfaces of revolution like the ellipsoid-cylinder and that the mean coefficient is referred to the ellipsoidal surface area, which is supposed to be isothermal.

Hemisphere-Cylinder ($L/a = 1$). Stalder and Nielsen (8) performed tests on a hemisphere-cylinder of 1 in. diam. The hemisphere was made of copper and, presumably, was at uniform temperature. They measured mean coefficients directly. Their tests covered ranges of Mach number from 0.12 to 5.04 and of Reynolds number from 65,000 to 600,000, based on diameter as characteristic length and on values of fluid properties after the normal shock.

Ellipsoid-Cylinder ($L/a = 3$). Lewis and Ruggeri (3) performed isothermal tests on an ellipsoid-cylinder of 20 in. diam. The model was electrically heated in sections so that stepwise mean values could be measured. A fairing behind the nosepiece may have influenced the results. Our own results (not presented) from tests with a rising step, located one diameter downstream of the ellipsoidal surface, indicated that its effect was to reduce the heat transfer because it retarded the flow over the surface. A similar effect may have been present in the previous tests. Transition and turbulent transfer occurred in all their tests, the critical local Reynolds number being about 2×10^6 at zero angle of attack. The value of $f(L/a)$ in the tabulation is based on a single test at 152 knots and a free-stream Reynolds number of 3×10^6, transition occurring at a profile distance of 15 in. In order to get the mean laminar value, the authors' measurements of the heat transfer were integrated over the ellipsoid area down to the distance of 15 in.; then the heat transfer from 15 to 33.4 in. was calculated under the assumption that the boundary layer was laminar to the end of the profile. For this purpose flat-plate approximations were employed; they appeared justifiable because beyond 15 in. the velocity roof was virtually uniform.

TABLE 2 COMPARISON WITH OTHER EXPERIMENTAL VALUES

$\frac{L}{a}$	$f(L/a)$	$\sqrt{\frac{S}{D}} \cdot f(L/a) = C$	$\frac{A}{\pi D^2} \cdot f(L/a)$	Reference
1	1.03	0.92	0.515	[8]
3	0.60	0.78	0.74	[3]
4	0.52	0.76	0.85	Present Work

The variation of $A/(\pi D^2) \cdot f(L/a)$ in Table 2 reflects the influence of surface area on the total heat transfer for surfaces of fixed diameter D. For $L/a \gg 4$, it is expected that values of $(S/D)^{1/2} f(L/a)$ would approach the flat-plate value of 0.664, barring any effects of transition or transverse curvature.

CONCLUSION

The mass heat-transfer analog was employed to obtain mean coefficients of laminar heat transfer on ellipsoidal surfaces of 4:1 axis ratio, the general method following that of reference (7). A boundary-layer calculation gave results 10 per cent lower than the experimental value. In view of the turbulent nature of the free jet and of the approximative nature of the calculations, the agreement was considered satisfactory.

It is recommended that until additional information is available the mean value of f (4) in Equation [17] be taken to be 0.52 in the range of Reynolds number (based on diameter) from 15,000 to 130,000.

It was shown that, when reproducible sublimation data can be attained, a significant reduction of the scattering can be achieved by allowing for the effects of evaporative cooling and aerodynamic heating.

The result of the present experimentation was compared with results on related surfaces. In general they are expressed by Equation [17] with $f(L/a)$ given in Table 2. The values given are probably sufficiently accurate for preliminary estimates.

The general trend of the function $f(L/a)$ is found consistent but additional information is needed to fix it in all details. For this purpose the analog may be used with considerable advantage over other methods.

BIBLIOGRAPHY

1 "Method for Calculation of Heat Transfer in Laminar Region of Air Flow Around Cylinders of Arbitrary Cross Section," by E. R. G. Eckert and J. N. B. Livingood, NACA TN 2733, 1952.

2 "Aircraft Windshield Heat and Mass Transfer," by M. Jakob, S. P. Kezios, A. Sinila, H. H. Sogin, and M. Spielman, AF Technical Report No. 6120, Part 5, 1952, pp. 417–432.

3 "Investigation of Heat Transfer from a Stationary and Rotating Ellipsoidal Forebody of Fineness Ratio 3," by J. P. Lewis and R. S. Ruggeri, NACA TN 3837, November, 1956.

4 "Pressure Distribution from Theoretical Approximations of the Flow Pattern," by J. S. McNown and E. Y. Hsu, Heat Transfer and Fluid Mechanics Institute, Berkeley, Calif., 1949, pp. 65–76.

5 "Heat Transfer at the Forward Stagnation Point of Blunt Bodies," by E. Reshotko and C. B. Cohen, NACA TN 3513, July, 1955.

6 "Cavitation and Pressure Distribution-Head Forms at Zero Angle of Yaw," by H. Rouse and J. S. McNown, State University of Iowa, Studies in Engineering, Bulletin No. 32, 1948.

7 "Sublimation From Disks to Air Streams Flowing Normal to Their Surfaces," by H. H. Sogin, Trans. ASME, vol. 80, 1958, pp. 61–69.

8 "Heat Transfer from a Hemisphere-Cylinder Equipped With Flow-Separation Spikes," by J. R. Stalder and H. V. Nielsen, NACA TN 3287, 1954.

9 "The Evaporation of Naphthalene in Dry Air and in Moist Coal Gas," by J. S. G. Thomas, *Journal of the Society of Chemical Industry*, vol. 35, 1916, pp. 506–513.

[6] The coefficient in the third column is equivalent to C in Equation [11].

Investigation of Burnout Heat Flux in Rectangular Channels at 2000 Psia[1]

By H. S. JACKET,[2] J. D. ROARTY,[2] and J. E. ZERBE,[2] PITTSBURGH, PA.

Burnout heat-flux data were obtained under conditions of approximately zero exit quality and bulk boiling at the exit of electrically heated test specimens. These specimens were long, narrow channels with various slot thicknesses, surfaces, materials, and length-to-diameter ratios. Tests were run at 2000 psia and mass velocities from approximately 0.2×10^6 to 3×10^6 lb/hr-sq ft. The effect of inclining the channel at 45 deg also was investigated. The rectangular channel burnout results are in reasonable agreement with data previously obtained for round tubes. A design equation is suggested which yields a conservative estimate of the burnout heat flux in the low subcooling and quality regions for the range of variables investigated herein. A burnout loop and method of operation are described.

Nomenclature

The following nomenclature is used in the paper:

D = equivalent diameter of test section, ft
G = mass velocity, lb/hr-sq ft
H = enthalpy of mixture, Btu/lb
L = test section heated length, ft
ϕ = heat flux density, Btu/hr-sq ft
rms = root mean square, microin.
x = bulk quality of liquid—vapor mixture, mass fraction of vapor
T = temperature of bulk liquid, deg F
ΔT = saturated temperature minus test water bulk temperature, deg F

Subscripts

BO = burnout
in = inlet to test section

Terminology

Local or Subcooled Boiling. This occurs with the surrounding liquid mostly at a temperature below saturation. The bubbles usually condense in the subcooled liquid and as a result, there is no net retention of vapor.

Saturated or Bulk Boiling. This occurs in a liquid where the temperature is equal to or slightly higher than saturation; it implies a net generation of vapor.

Nucleate Boiling. Vapor is formed as discrete bubbles; a condition which is characteristic of wetted heating surfaces.

Film Boiling. Condition developed wherein vapor exists as a continuous film on the heating surface.

Burnout Heat Flux. Maximum heat flux under nucleate boiling conditions before vapor blanketing begins.

[1] This work was done under U. S. Atomic Energy Commission Contract AT-11-1-GEN-14.

[2] Bettis Atomic Power Division, Westinghouse Electric Corporation.

Contributed by the Heat Transfer Division and presented at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, November 16, 1956. Paper No. 57—SA-6.

Introduction

A knowledge of the heat flux under which physical burnout of a heat-transfer surface occurs is of prime importance for boiling systems with forced circulation of the coolant. Water-cooled and/or moderated nuclear reactors must be designed to avoid physical burnout under the various abnormal conditions that might occur during operation. Information on maximum heat flux or burnout flux is, therefore, essential to this type of reactor-core design.

A program of burnout testing has been conducted to investigate the effects of (1) geometry, (2) length-to-diameter ratio, (3) test-section orientation, (4) power pulse, and (5) surface material and finish on burnout heat flux in both the local and bulk-boiling regions. The purpose of this investigation is to provide the necessary data for reactor-core designs where boiling and high heat-flux conditions might exist during operation of the reactor.

Apparatus and Test Procedure

The burnout loop was designed with two test legs to accommodate both burnout and pressure-drop experimentation at pressures up to 2000 psia. Both test sections are heated by a 480-kw d-c power supply. The maximum flow to either section is 30 gpm. A schematic diagram of the facilities is shown in Fig. 1. The parallel channel leg shown in the figure was not used for the tests discussed here. The major loop components and instrumentation are described in the Appendix.

Test Specimens. An exploded view of a typical test-specimen assembly is shown in Fig. 2. To obtain closer simulation of an actual reactor channel, the test specimens were designed to burn out on the flat plate rather than at the corners by reducing the thickness of the cross section in each corner. This reduction was appropriately accounted for in calculations of the heat-flux values. In the direction of flow, the heat flux was uniform; thus, burnout always occurred at the downstream end of the channel.

To determine the effect of different channel lengths and equivalent diameters, materials, specimen construction, and specimen surfaces on the burnout flux, the following variables were built into the specimens:

Channel material: Commercial grade "A" nickel and Zircaloy-2
Channel flow spacing: 0.050 in., 0.055 in., and 0.097 in.
Channel width: 1 in.; heated width: 0.88 in.
Channel length: $12^1/_{16}$ in. and 27 in.
Channel manufacturing technique: Roll-bonded or welded
Heat transfer surface: Machined (32 microin. rms roughness) or hot rolled and pickled (140 microin. rms roughness)

Test Procedure. The loop was filled with cold, demineralized water and degassed for 4 to 8 hr. Flow was then passed through the ion exchanger until a minimum purity of 2 megohm-cm was obtained. The loop was raised to the desired test pressure and the preheaters set to give required temperature conditions. The flow rate was set to the desired value and power applied to the test section at a rate that permitted all conditions to remain essentially in thermal equilibrium throughout the test. The burnout point was determined when the burnout detector tripped the circuit breaker of the power supply or when the exit-wall thermo-

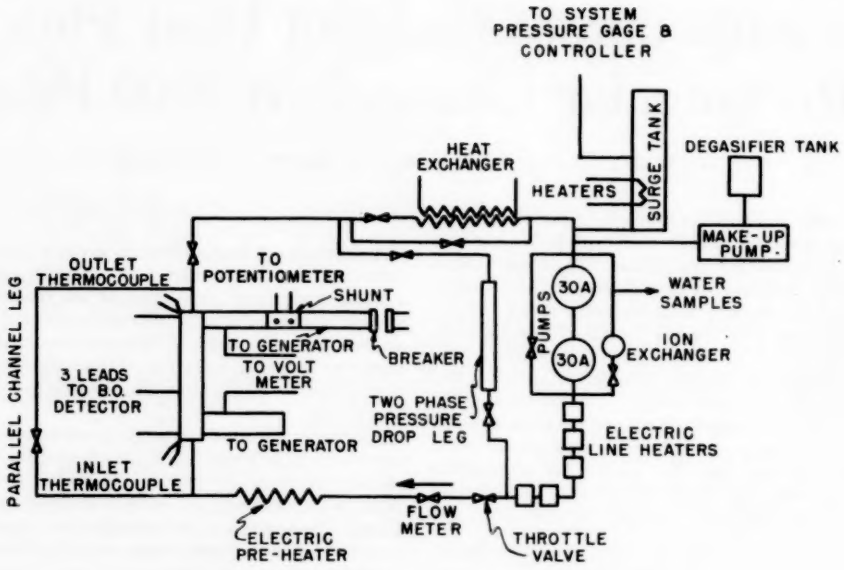


FIG. 1 SCHEMATIC DIAGRAM OF BURNOUT LOOP

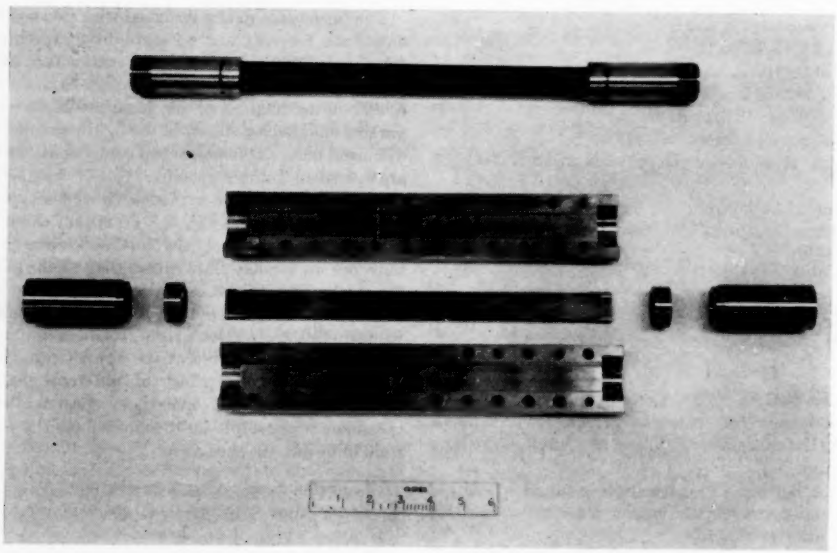


FIG. 2 TYPICAL TEST-SPECIMEN ASSEMBLY

couple exhibited an excursion or a steadily increasing reading without an accompanying increase in test-section power. The temperature excursion was generally greater than 50 F.

The test conditions were as follows:

1 *Approximately Zero Exit Subcooling (Zero Exit Quality) Burnout Tests*

(*a*) 2000 psia, vertical, upflow.

(*b*) Mass velocity range: 0.2×10^6 to 1.0×10^6 lb/hr-sq ft.

2 *Quality Burnout Tests*

(*a*) 2000 psia, vertical and inclined 45 deg, upflow.

(*b*) Mass-velocity range: 0.15×10^6 to 3.0×10^6 lb/hr-sq ft.

(*c*) Exit steam-quality range: 0 to 100 per cent.

ACCURACY OF RESULTS

The method of Kline and McClintock[3] was used to evaluate the uncertainty interval of the burnout-flux results due to uncertainties in each variable involved in the computation.

The reported burnout heat fluxes are average values over the heated specimen and are based on power and heat-transfer-area

[3] "Describing Uncertainties in Single-Sample Experiments," by S. J. Kline and F. A. McClintock, *Mechanical Engineering*, vol. 75, 1953, pp. 3–8.

TABLE 1 UNCERTAINTY OF VARIABLES

Voltage	$v:w_v$	±0.005 v
Current	$I:w_I$	±0.0016 I
Flow Rate, gpm	$W:w_W$	±0.005 W
Test Section Length	$L:w_L$	±1/64 in.
Channel Spacing	$D:w_D$	±0.006 in. for 0.055-in. channels ±0.004 in. for 0.097-in. channels
Inlet Temperature*	$T_i:w_{T_i}$	±2F°
Outlet Temperature*	$T_o:w_{T_o}$	±2F°
Fluid Density	$\rho:w_\rho$	±0.02ρ
Channel Width	$b:w_b$	±0.001 in.

* The inlet and exit temperatures are used to calculate the enthalpies of the fluid.

measurements with an error of less than ±1 per cent for 20:1 odds (95 per cent confidence in the specified error of the individual measurements). Considering a possible maximum discrepancy in a heat balance of 5 per cent as discussed later, the uncertainty in the burnout flux would be approximately +1 and −6 per cent.

A heat balance was made to determine the agreement between the amount of heat picked up in the water and the power dissipated in the test section. In the subcooled region, the heat balance generally checked to within 5 per cent. In the quality region, it was not possible to make a heat balance because of the lack of an independent measurement of the exit steam quality.

In the subcooled region, the interval of uncertainty in the calculation (for 95 per cent confidence limit) of the heat picked up in the water was found to be approximately ±2.5 per cent. This error is based on a water-temperature rise of 200 F or a corresponding enthalpy rise (a mean value for the results reported here) as the water passes through the test section.

The difference between the discrepancy in the heat balance and the estimated error caused by errors in the measurement taken is probably due to heat leakage from the test specimen to its environment and to its terminals. It is believed that the accuracy of the heat balance in the quality region is approximately the same as in the subcooled region.

In the course of the investigations, the test specimens had a tendency to expand against the backup housing and thus increase the nominal channel spacing. Because of this increase, the inaccuracy of the channel dimensions is rather large, as shown in Table 1. The interval of uncertainty (for 95 per cent confidence limit) in the calculation of G, the mass velocity, is ±11.0 per cent for the 0.055-in. channels and ±4.6 per cent for the 0.097-in. channels. The weight flow, which is independent of the spacing, is accurate to within ±2.1 per cent.

The interval of uncertainty associated with each variable employed in these calculations is estimated for 95 per cent confidence level as shown in Table 1.

RESULTS AND DISCUSSION

General. Tests were run to determine the effects on burnout of (1) geometry, (2) length-to-diameter ratio, (3) test-section orientation, (4) power pulse, and (5) surface material and finish. Length-to-diameter ratios of 64, 120, and 140 were investigated. The data for approximately zero exit quality are given in Table 2

TABLE 2 ZERO EXIT QUALITY AND SUBCOOLED BURNOUT DATA; EXIT PRESSURE 2000 PSIA

Test Channel	Exit Subcooling F°	Mass Velocity 10^6 lb/hr-ft²	Burnout Flux* 10^6 Btu/hr-ft²
Vertical nickel (machined)	3	1.35	1.120
Flow dimension: 1 in. x 0.055 in.	5	1.32	0.920
Length: 12-1/16 in.	2	0.926	0.938
	4	0.667	0.949
	6	0.578	0.786
	3	0.269	0.360
Vertical Zircaloy-2 (hot-rolled)	3	1.47	1.42
Flow dimension: 1 in. x 0.055 in.	12	0.919	0.765
Length: 12-1/16 in.	11	0.858	0.786
	4	0.571	0.695
	0	0.195	0.300**
Vertical Zircaloy-2 (machined)	11	1.24	0.979
Flow dimension: 1 in. x 0.097 in.	14	0.888	0.921
Length: 12-1/16 in.	3	0.529	0.851
	4	0.208	0.333
Vertical Zircaloy-2 (machined)	3	1.21	1.11
Flow dimension: 1 in. x 0.097 in.	4	0.906	0.847
Length: 27 in.	4	0.906	0.841
	4	0.604	0.747

* Burnout Flux is evaluated from 94% of the total power as generated in the thick portion of the rectangular channel (6% of the power is generated in the thin edges). The surface area of the thick portion of the channel is 0.88 in. x length x 2.

** This value is estimated as the maximum heat flux in the nucleate boiling region. Physical burnout occurred in the film boiling region at a heat flux of approximately 0.560 x 10^6 Btu/hr-ft².

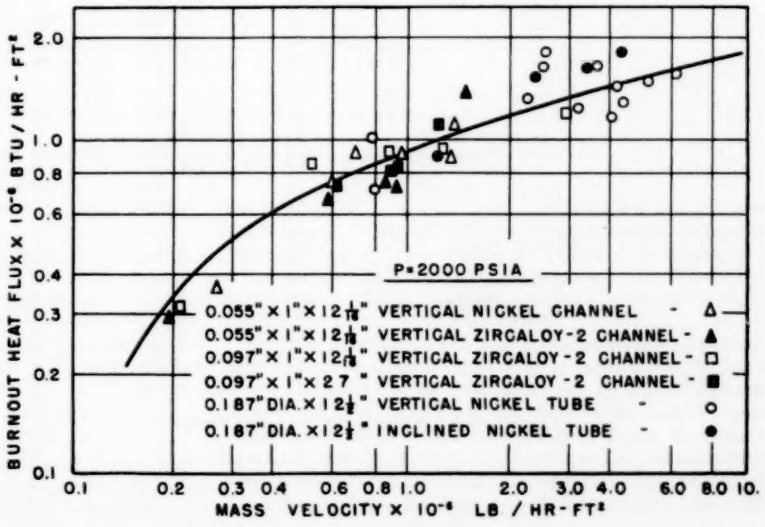


FIG. 3 BURNOUT HEAT FLUX VERSUS MASS VELOCITY FOR APPROXIMATELY ZERO EXIT QUALITY AND ZERO EXIT SUBCOOLING

TABLE 3 QUALITY BURNOUT DATA, EXIT PRESSURE 2000 PSIA

Test channel	Inlet subcooling, deg F	Mass velocity, 10^6 lb/hr-sq ft	Burnout flux, 10^6 Btu/hr-sq ft	Quality, per cent
Vertical Zircaloy-2 (machined)	8	2.67	0.808	12.3
	11	2.57	0.860	13.0
Flow dimension: 1 in.	8	1.94	0.771	17.0
× 0.097 in	9	1.38	0.718	22.6
Length: $12^1/_{16}$ in.	8	1.19	0.773	29.2
	9	0.864	0.704	37.1
	10	0.479	0.558	54.3
	10	0.187	0.376	97.0
	36	2.74	1.08	7.7
	36	2.74	1.10	8.1
	34	2.72	1.08	8.4
	37	2.10	0.970	11.0
	36	1.28	0.839	20.5
	63	2.92	1.24	0.9
	60	2.17	1.08	5.7
	62	1.41	0.969	14.0
	60	1.32	0.959	16.9
	60	0.924	0.760	21.6
	60	0.466	0.618	46.7
	62	0.179	0.402	92.5
Inclined (45 deg) Zircaloy-2 (machined)	11	1.22	0.731	25.8
	12	0.843	0.625	32.4
Flow dimension: 1 in.	11	0.443	0.502	52.6
× 0.097 in.	10	0.171	0.330	93.3
Length: $12^1/_{16}$ in.	58	1.31	0.974	17.5
	59	0.915	0.823	25.2
	60	0.469	0.622	46.3
	60	0.176	0.375	87.2
Vertical Zircaloy-2 (machined)	10	2.71	0.557	19.2
	10	2.03	0.504	24.1
Flow dimension: 1 in.	10	1.35	0.448	33.0
× 0.097 in.	9	1.13	0.436	39.6
Length: 27 in.	10	0.840	0.397	48.7
	10	0.438	0.320	76.9
	9	0.157	0.170	100.0[a]
	63	3.29	0.902	10.7
	62	2.88	0.848	12.2
	61	2.24	0.732	15.7
	66	1.49	0.594	24.6
	62	1.32	0.565	27.4
	62	0.937	0.487	37.4
	62	0.467	0.392	72.6
	62	0.187	0.207	100.0[a]
Inclined (45 deg) Zircaloy-2 (hot rolled)	11	1.33	0.596	35.4
	9	0.894	0.496	45.6
Flow dimension 1 in.	9	0.473	0.403	72.0
× 0.055 in.	10	0.175	0.242	100.0[a]
Length: $12^1/_{16}$ in.	60	1.42	0.755	23.5
	59	0.955	0.613	37.6
	60	0.499	0.472	64.0
	60	0.173	0.253	100.0[a]
Inclined (45 deg) Zircaloy-2 (hot-rolled)	35	1.47	0.700	34.0
	34	1.46	0.697	34.7
Thin edge up	36	1.10	0.558	36.1
Flow dimension: 1 in.	34	1.03	0.598	44.9
× 0.050 in.	37	1.02	0.602	44.2
Length: $12^1/_{16}$ in.	38	0.494	0.446	74.6
	37	0.490	0.446	75.6
Vertical Zircaloy-2 (hot-rolled)	335	1.05	1.142	18.55
	335	0.626	0.842	43.2
Flow dimension: 1 in.	335	0.333	0.573	79.5
× 0.050 in.	136	1.91	0.855	5.6
Length: $12^1/_{16}$ in.	137	0.891	0.713	36.7
	135	0.595	0.595	56.6
	11	1.58	0.586	30.6
	11	0.842	0.526	55.9
	11	0.578	0.509	81.6

[a] Exit fluid is actually a few degrees superheated.

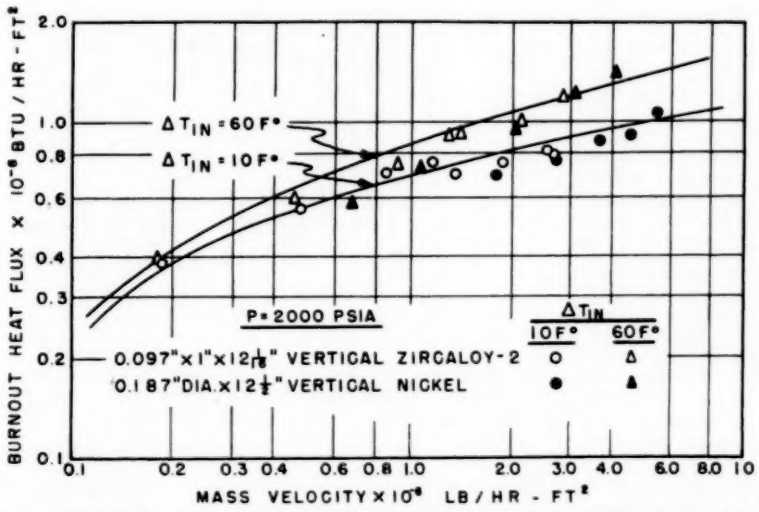


FIG. 4 BULK BOILING BURNOUT HEAT FLUX VERSUS MASS VELOCITY FOR 10 F AND 60 F INLET SUBCOOLINGS

(NOTE: The enthalpy at exit of the round tube and rectangular channel is approximately equal for a given heat flux, mass velocity, and inlet subcooling.)

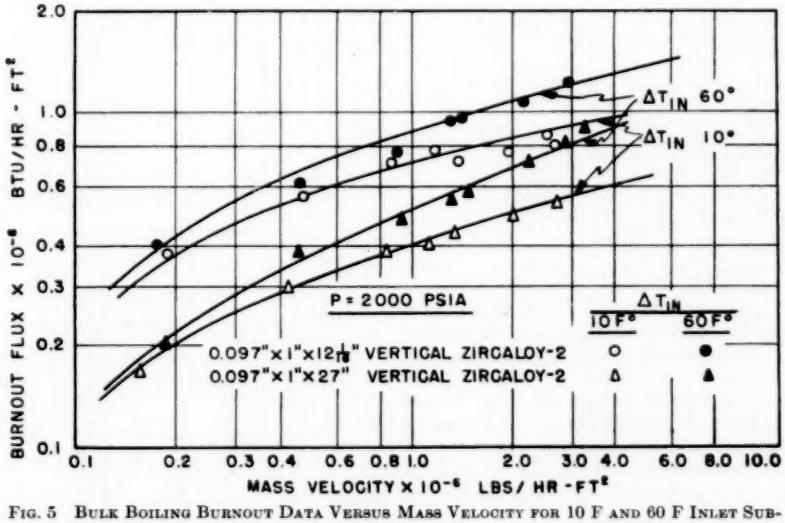


FIG. 5 BULK BOILING BURNOUT DATA VERSUS MASS VELOCITY FOR 10 F AND 60 F INLET SUBCOOLINGS

(NOTE: The enthalpy at the exit of each channel is different for a given heat flux, mass velocity, and inlet subcooling.)

and plotted in Fig. 3. The quality or bulk-boiling burnout-heat-flux data are given in Table 3 and are plotted in Figs. 4 and 5.

Effect of Geometry. In the early test operation of the burnout loop, round tube-test sections were used to check existing local boiling-burnout data.[4] Some of the round-tube data for approximately zero quality and also for bulk boiling are shown in Figs. 3 and 4. These round-tube data appear to give the same values of burnout heat flux as data for rectangular channels having channel thicknesses of 0.055 and 0.097 in.

Effect of L/D. The *L/D* effect is interpreted to mean any variation in the burnout flux due to changing the heated length of the channel while the local fluid conditions at the burnout point are kept constant. This series of tests was run on machined Zircaloy-2 test specimens $12^1/_{16}$ in. and 27 in. long (L/D = 64 and 140). In the low-subcooled region, Fig. 3, a variation in length-to-diameter ratio between 64 and 140 appears to have no effect on burnout heat flux.

Data were taken on a $12^1/_{16}$-in-long channel at an inlet subcooling of 11 F and a mass velocity of 0.2×10^6 to 2.5×10^6 lb/hr-sq ft. A comparative set of data was run on a 27-in. channel; however, the inlet temperature was adjusted so that 11 F subcooling existed $12^1/_{16}$ in. from the exit of the channel. In conducting the test this way, the *L/D* effect attributable to a heat balance is separated from any additional *L/D* effect that might be present. If any discrepancy existed between the data obtained from the

[4] "Analysis of Heat Transfer, Burnout, Pressure Drop and Density Data for High Pressure Water," by W. H. Jens and P. A. Lottes, AEC Research and Development Report ANL 4627, May 1, 1951.

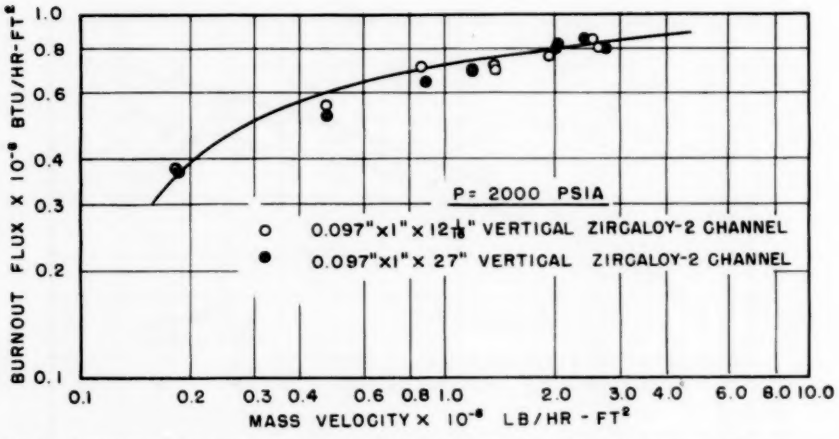


FIG. 6 INVESTIGATION OF THE EFFECT OF L/D RATIO ON BURNOUT FLUX FOR VERTICAL RECTANGULAR CHANNELS

(NOTE: The enthalpy of the fluid at the exit was the same in each channel for a given heat flux and mass velocity.)

TABLE 4 COMPARISON OF BURNOUT FLUX (SUBCOOLED AND BULK BOILING) IN ZIRCALOY-2 CHANNELS (MACHINED); EXIT PRESSURE 2000 PSIA

Mass Velocity 10^6 lb/hr-ft²		Burnout Flux 10^6 Btu/hr-ft²		Quality %	
12-1/16 in.	27 in.	12-1/16 in.	27 in.	12-1/16 in.	27 in.
2.67	2.74	0.808	0.892	12.3	11.8
2.57	2.47	0.860	0.806	13.0	13.0
1.94	2.03	0.771	0.793	17.0	14.3
	2.08		0.806		13.4
1.38	1.38	0.718	0.712	22.6	21.9
1.19	1.21	0.773	0.699	29.2	24.6
0.864	0.897	0.704	0.649	37.1	29.5
0.479	0.481	0.588	0.522	54.3	49.0
0.187	0.188	0.376	0.367	97.0	88.3
				Exit Subcooling, F°	
1.24	1.21	0.979	1.11	11	3
0.888	0.906	0.921	0.847	14	4
	0.906		0.841		4
0.529	0.605	0.851	0.747	3	4

two channels, information on the degree of mixing or validity of calculating the exit quality from the first law could be derived. Data from these runs are reported in Table 4 and compared in Fig. 6.

Fig. 6 indicates that the burnout flux for the 27-in. channel (L/D = 140) is approximately equal to the corresponding flux for a $12^1/_{16}$-in. channel. This evidence may be interpreted to mean that there is excellent mixing in the channels and that the calculated exit quality may be a significant parameter for predicting burnout.

In the L/D effect tests described, it was necessary to increase the inlet subcooling with the longer channel in order to obtain the same bulk coolant conditions at burnout at the end of the channel. It is thus possible that two opposing effects are occurring which make the burnout fluxes approximately equal; namely, the effect of L/D and the effect of inlet subcooling. The comparison in Table 5 indicates the possible effect of inlet subcooling on burnout for a given geometry (27 in. long, 97-mil channel); in general, the higher the inlet subcooling, the higher the burnout flux. For this comparison the following conditions are maintained:

(*a*) Approximately constant exit bulk quality.

(*b*) Constant mass velocity or larger mass velocity for the condition yielding the lower value of burnout flux.

Another method of comparison is shown in Table 6. In this table a comparison of burnout points for 97-mil channels, $12^1/_{16}$ in. and 27 in. long, was made on the following basis:

TABLE 5 POSSIBLE EFFECT OF INLET SUBCOOLING ON BURNOUT

Comparison Set	Quality %	Mass Velocity 10^6 lb/hr-ft²	ΔT_{IN} F°	Burnout Flux 10^6 Btu/hr-ft²
I	88.3	0.188	513	0.367*
I	100.0	0.187	62	0.207
II	49.0	0.481	263	0.522
II	48.5	0.840	10	0.397
III	29.5	0.897	177	0.649
III	27.4	1.32	62	0.565**
IV	24.6	1.21	133	0.699
IV	24.6	1.49	66	0.594
IV	24.1	2.03	10	0.504
V	13.0	2.47	73	0.806
V	14.3	2.03	94	0.793
V	15.7	2.24	61	0.732

* For comparison purposes, if quality = 100%, burnout flux may be estimated as approximately 0.325 x 10^6.

** For comparison purposes, if quality = 29.5%, burnout flux may be estimated as approximately 0.550 x 10^6.

TABLE 6 POSSIBLE EFFECT OF L/D ON BURNOUT

Comparison Set	ΔT_{IN} F°	Quality %	Mass Velocity 10^6 lb/hr-ft²	Length in.	Burnout Flux 10^6 Btu/hr-ft²
I	62	14.0	1.41	12-1/16	0.969
I	62	12.2	2.88	27	0.848
II	60	16.9	1.32	12-1/16	0.959
II	61	15.7	2.24	27	0.732
III	60	21.6	0.924	12-1/16	0.760
III	66	24.6	1.49	27	0.594
IV	60	46.7	0.466	12-1/16	0.618
IV	62	37.4	0.937	27	0.487*
V	62	92.5	0.179	12-1/16	0.402
V	62	72.6	0.467	27	0.392**
VI	8	29.2	1.19	12-1/16	0.773
VI	10	24.1	2.03	27	0.504
VII	9	37.1	0.864	12-1/16	0.704
VII	10	33.0	1.35	27	0.448
VIII	10	97.0	0.187	12-1/16	0.376***
VIII	9	100.0	0.157	27	0.170

* For comparison purposes, if quality = 46.7%, burnout flux may be estimated as approximately 0.450 x 10^6.

** For comparison purposes, if quality = 92.5%, burnout flux may be estimated as approximately 0.300 x 10^6.

*** For comparison purposes, if mass velocity = 0.157 x 10^6, burnout flux may be estimated as approximately 0.300 x 10^6.

(*a*) Constant inlet subcooling.

(*b*) Approximately constant exit bulk quality.

(*c*) Mass velocity greater for the condition that yields the lower value of burnout flux.

From Table 6 it is noted that, although the inlet temperature and exit quality are the same for both channels and the mass velocity is higher for the longer channel, burnout in general occurs at a lower heat flux in the longer channel.

TABLE 7 SUBCOOLED AND BULK BOILING BURNOUT FLUX IN ZIRCALOY-2 HOT-ROLLED CHANNELS INCLINED 45 DEG, EXIT PRESSURE 2000 PSIA

(Channel flow dimensions: 0.050 in. by 1 in. Length: $12^{1}/_{16}$ in.)

Inlet temp, deg F	Inlet velocity, fps	Mass velocity, G 10^6 lb/hr-sq ft	Exit subcooling or quality, deg F or per cent	Burnout flux ϕ_{BO}, 10^6 Btu/hr-sq ft	Remarks
301	2.08	0.431	49.6%	0.611	Burnout indicated by excessive wall temperature.
300	2.06	0.428	53.3%	0.642	Previous run reproduced.
300	2.06	0.428	52.8%	0.622	Heat flux applied in increments of 10,000 Btu/hr-sq ft every 2 min, commencing well below ϕ_{BO} and increasing until burnout occurred.
301	2.00	0.417	54.3%	0.611	Heat flux applied in increments of 10,000 Btu/hr-sq ft every min, commencing well below ϕ_{BO} and increasing until burnout occurred.
301	2.08	0.431	55.6%	0.638	Same as above.
302	5.03	1.04	1 F	0.909	Burnout indicated by wall-temp excursion.
300	5.09	1.05	1 F	0.924	Instability occurred indicating incipient burnout.
300	5.08	1.05	1 F	0.924	Previous instability reproduced.
300	5.06	1.05	11.4%	1.05	Heat flux applied in increments of 5000 Btu/hr-sq ft every 2 min, commencing at $\phi = 0.820 \times 10^6$, until instability was noted at $\phi = 0.911 \times 10^6$. Flux was increased until wall-temp excursion occurred at $\phi_{BO} = 1.05 \times 10^6$.
301	5.16	1.07	2 F	0.939	Previous run reproduced with different test section and operating crew.
296	9.00	1.85	31 F	1.44	Burnout indicated by excessive wall temp.
500	5.8	1.03	15.26%	0.581	Instability first noted at $\phi = 0.528 \times 10^6$ Btu/hr-sq ft at $G = 1.03 \times 10^6$ and quality = 11.6 per cent. Heat flux was increased until burnout occurred at ϕ_{BO} = 0.581 Btu/hr-sq ft.
501	5.4	0.951	19.7%	0.562	Heat flux applied in increments of 20,000 Btu/hr sq ft every min, commencing at $\phi = 0.350 \times 10^6$ Btu/hr-sq ft. Instability observed at $\phi = 0.436 \times 10^6$ Btu/hr-sq ft, $G = 0.957 \times 10^6$ and quality = 7 per cent. Heat flux was increased until burnout occurred.
500	6.1	1.09	31.7%	0.798	Burnout indicated by excessive wall temp.
501	6.2	1.10	32.7%	0.810	Previous run reproduced.
500	9.8	1.74	2.95%	0.722	Instability occurred indicating incipient burnout.
500	9.5	1.69	7.4%	0.770	Heat flux applied in increments of 20,000 Btu/hr-sq ft every min, commencing at a flux well below burnout. Instability noted at $\phi = 0.708 \times 10^6$ Btu/hr-sq ft, $G = 1.69 \times 10^6$ and quality = 3.7 per cent. Heat flux was increased until burnout occurred.
599	5.4	0.828	45.15%	0.495	Incipient burnout.
600	5.2	0.799	46.9%	0.492	Heat flux applied in increments of 10,000 Btu/hr-sq ft every min, commencing at $\phi = 0.302 \times 10^6$ Btu/hr-sq ft. Incipient burnout was indicated at $\phi_{BO} = 0.492 \times 10^6$ by instability in wall temp.
600	5.0	0.800	45.99%	0.477	Heat flux applied in increments of 20,000 Btu/hr-sq ft every min, commencing at $\phi = 0.315 \times 10^6$. Incipient burnout was indicated by severe instability in wall temp.
599	9.7	1.5	29.7%	0.660	Burnout indicated by excessive wall temp.

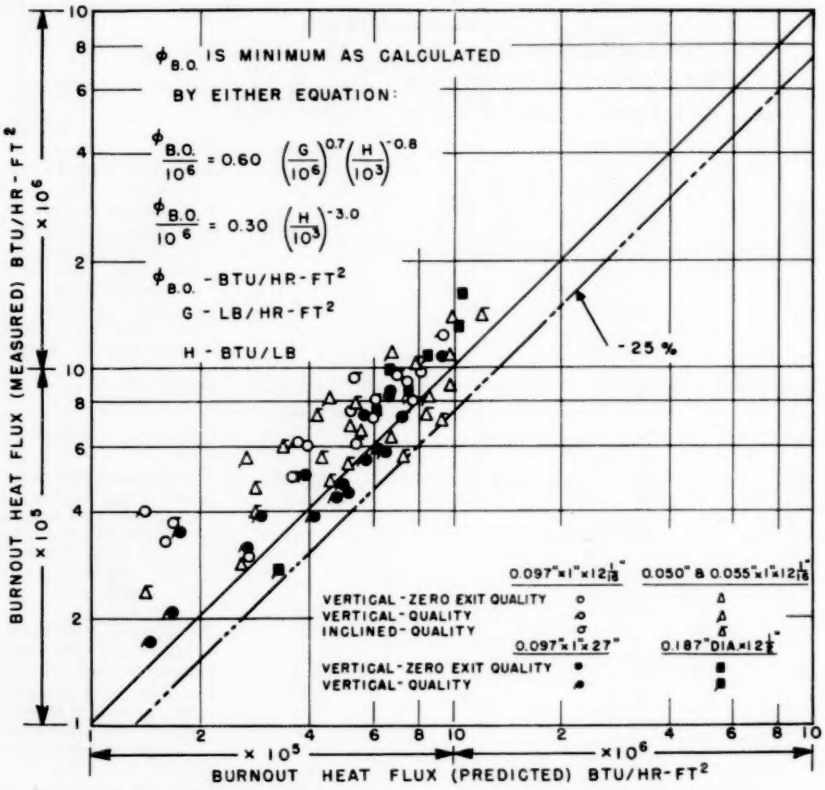


FIG. 7 COMPARISON OF BETTIS DATA WITH BURNOUT DESIGN EQUATION (P = 2000 PSIA)

Effect of Inclination. Burnout tests were run on 97-mil smooth, and 50 and 55-mil hot-rolled Zircaloy specimens inclined at an angle of 45 deg with the thin edge in the vertical and horizontal positions. The data are tabulated in Tables 3 and 7. All inclined specimens were observed to burnout on the upper side only, an indication that preferential stratification did occur at the burnout point.

Zero exit subcooling data for a vertical and an inclined 187-mil round tube are compared in Fig. 3. No effect on burnout due to inclination can be seen.

Quality burnout data were obtained for a 50-mil channel with the thin edge in the vertical position. These data are reported in Table 3. It appears that this orientation of the test section has no significant effect on burnout heat flux. In Fig. 7, vertical and inclined 50, 55, and 97-mil data (with thin edge horizontal) are compared with a suggested design equation. The 50-mil inclined quality data are somewhat lower than expected. This is probably due to instabilities which were noticed.

Other investigators[5] have reported a sudden instability in a test system causing premature burnouts at fluxes well below the expected burnout flux. Similar instabilities have been encountered occasionally while obtaining the data reported here. The results of specific tests to investigate these instabilities are reported in Table 7 and may be summarized as follows:

(*a*) Instabilities occurred at heat fluxes as much as 21 per cent below the actual burnout heat flux. Such instabilities may be partially responsible for lower burnout in the case of inclined 50-mil channels.

(*b*) Neither the exact nature nor the cause of the instabilities was determined. Loop effects such as control-valve chatter and lack of sufficient pump head were investigated, and appeared to have little effect on the instability. The rate of heat flux application did not affect burnout in the ranges investigated.

Particular attention was devoted to ascertaining whether or not an instability in flow or autocatalytic effect as discussed by Jens[6] accompanied burnout. This effect was not detected. The flow meter used in these investigations indicated no appreciable decrease in flow at the burnout point.

Effect of Power Pulse on Burnout. Preliminary power-pulse tests are reported in Table 8. Pulsing the power from about 20 per cent below the burnout point to the burnout point apparently had no effect on burnout for the few tests conducted. The rate of heat-flux application was approximately 100,000 to 200,000 Btu/hr-sq ft per sec.

Effect of Surface and Material. Fig. 3 includes a comparison of the machined Zircaloy-2 surface and the hot-rolled, pickled Zircaloy-2 surface. Despite the rather obvious difference in surface roughness (32 microin. rms for the machined surface compared to 140 microin. rms for the hot-rolled surface), no apparent effect on burnout was observed. For the few data available from this in-

[5] "Boiling Burnout Newsletter No. 2," by W. M. Rohsenow and J. A. Clark, Brookhaven National Laboratory BNL-2141, January 5, 1955.

[6] "Local Boiling Heat Transfer to Water at Low Reynolds Numbers and High Pressures," by J. A. Clark and W. M. Rohsenow, Trans. ASME, vol. 76, 1954, pp. 553–562.

TABLE 8 EFFECT OF POWER PULSE ON BURNOUT IN ZIRCALOY-2 HOT-ROLLED CHANNELS INCLINED 45 DEG; EXIT PRESSURE 2000 PSIA*

(Channel flow dimensions: 0.050 in. × 1 in.; length, 12 1/16 in.)

Inlet Temp °F	Inlet Vel fps	Mass Velocity 10^6 lb/hr-ft^2	Heat Flux 10^6 Btu/hr-ft^2 Initiation of Pulse	Heat Flux 10^6 Btu/hr-ft^2 Termination of Pulse**	Duration of Transient sec	Expected Steady-State Burnout Flux 10^6 Btu/hr-ft^2
301	5.0	1.04	0.692	0.853	1.5	0.900
300	5.0	1.04	0.692	0.858	1.5	0.900
300	5.0	1.04	0.692	1.31	2.5–3	0.900

* Exit subcooling not known because, due to the nonequilibrium condition of the loop, the exit water conditions were difficult to estimate.

** No burnout occurred in any of the tests.

vestigation there was also no observable difference in results obtained on a nickel specimen and on a Zircaloy-2 specimen.

CORRELATION OF RESULTS

The burnout data obtained in the series of experiments conducted cannot be applied directly to nuclear-reactor design because the reactor heat-flux distribution in the direction of flow is not uniform as was that of the cases tested herein. To utilize the burnout data in reactor design, it is desirable that burnout be correlated on the basis of the local fluid conditions. An expression was developed which is believed to give approximations of burnout heat flux within the ranges of variables investigated. No generality is intended in this expression, but it is useful as a design equation provided that (*a*) no extrapolation of the equation to regions beyond the limits of the variables investigated is made except for preliminary evaluations and (*b*) the value calculated by Equations [1] and [2] is reduced by 25 per cent.

Burnout heat flux is defined as the minimum value calculated from either of the following two equations

$$\frac{\phi_{BO}}{10^6} = 0.60 \left(\frac{G}{10^6}\right)^{0.7} \left(\frac{H}{10^3}\right)^{-0.8} \quad \text{.....[1]}$$

$$\frac{\phi_{BO}}{10^6} = 0.30 \left(\frac{H}{10^3}\right)^{-3.0} \quad \text{.....[2]}$$

where

ϕ_{BO} = burnout flux, Btu/hr-sq ft
G = mass velocity, lb/hr-sq ft
H = enthalpy of fluid, Btu/lb

Limits on the equation are as follows:

Pressure: 2000 psia
Geometry: Round tubes and rectangular channels similar to those tested
Mass velocity: 0.2×10^6 lb/hr-sq ft to 4×10^6 lb/hr-sq ft
L/D = 60 to 140
Local enthalpy at burnout point: 650 to 1135 Btu/lb
Test section position: 45 deg and vertical

The possible influences of the past history effects (inlet temperature and L/D) on burnout were considered in establishing Equations [1] and [2]. These equations were developed so that a conservative estimate of burnout heat flux is given; namely, the case when water at the saturation temperature is supplied at the inlet to the longest test channel investigated. Under these circumstances any inlet subcooling effect tends to increase the burnout heat flux. The L/D limit of 140 is very important because increasing L/D appears to lower the burnout flux.

In Fig. 7, Equations [1] and [2] are compared with the Bettis Plant burnout data.

Fig. 8 is a graph of Equations [1] and [2]. On this graph burnout appears to be a function of both mass velocity and bulk quality or enthalpy in certain regions. However, there are also regions where burnout appears dependent only on the fluid enthalpy or, in other words, the effect of mass velocity is insignificant.

CONCLUSIONS

The following conclusions can be drawn:

1 Zero exit subcooling, i.e., zero exit quality, burnout flux increases as the mass velocity increases in the range from approximately 200,000 to 1,000,000 lb/hr-sq ft. Few data are available at higher mass velocities, but, based on the behavior of low-quality

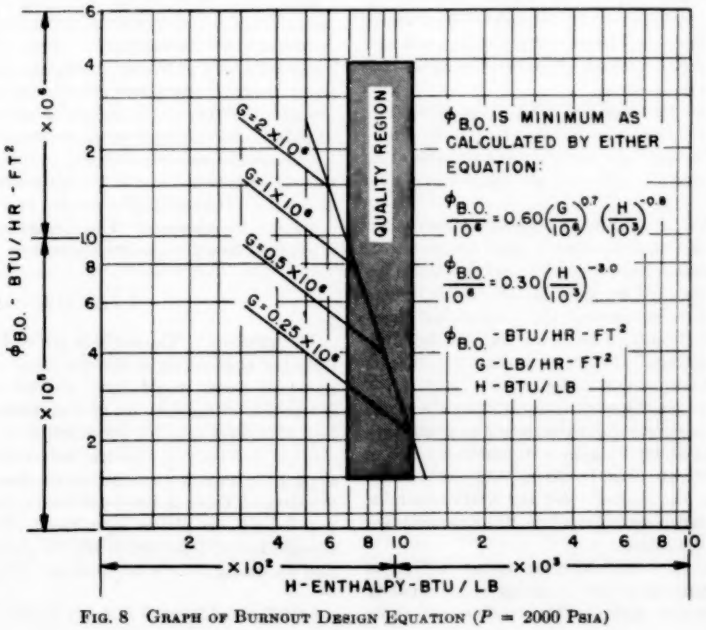


FIG. 8 GRAPH OF BURNOUT DESIGN EQUATION (P = 2000 PSIA)

burnout data in the range of 1,000,000 to 4,000,000 lb/hr-sq ft, it appears that zero exit quality and subcooling burnout flux increase only slightly for mass velocities above G = 1,000,000 lb/hr-sq ft.

2 There is no significant effect on burnout flux for conditions of approximately zero exit quality and zero exit subcooling as a result of changing channel-slot thickness (from 55 to 97 mil), channel geometry (round tube or narrow rectangular channels of approximately the same equivalent diameter), channel length, heat-transfer surface finish, or heat-transfer material.

3 Quality burnout results, as compared with a suggested design equation are essentially independent of channel position (vertical or 45 deg), geometry (round tubes or rectangular channels), heat-transfer surface finish, or heat-transfer materials.

4 In the bulk boiling region, there are apparent inlet subcooling and length-to-diameter ratio effects on burnout. For approximately constant local fluid conditions, increasing the inlet subcooling tends to increase the burnout heat flux, whereas increasing the L/D ratio tends to decrease the burnout heat flux.

Acknowledgments

The authors wish to express their gratitude to the Westinghouse Electric Corporation and to the U. S. Atomic Energy Commission for permission to publish this work. An expression of gratitude is also appropriate to Messrs. S. J. Green, R. A. DeBortoli, A. Weiss, T. W. Hunt, and S. W. Cota for assistance in the conduct of the investigation and in the design and operation of the laboratory equipment Assistance by the Bettis Information and Publications Group in the preparation of the manuscript is also acknowledged.

Appendix

Loop Components

Power Supply. Power is supplied by a 480-kw d-c generator designed to operate at 12,000 amp and 40 volts. However, for short periods, as much as 20,000 amp can be drawn from the generator.

Piping System. The main piping system is constructed of 1 1/2-in., schedule-80, type 347 stainless steel. All main system valves have type 316 cast stainless-steel bodies and stellite seats and plugs. The valves are Teflon packed. The throttle valves are globe type; all others are gate valves.

Pumps. Two Westinghouse oil-cooled, 30-A sealed pumps are used to circulate the water. The pumps may be run individually or connected in series. The control circuits are so interlocked that the pumps must be in operation before the generator can deliver power to the test section.

Pressurizing Tank. This vessel is used both to pressurize and degas the system. Pressurizing is accomplished by means of eight heaters which provide a total of 30 kw; six are manually controlled and two are operated by an automatic pressure controller. A liquid-level controller operates the system's high-pressure make-up pump. If the liquid level becomes too low, the main circulating pumps are shut off automatically. Shutting off these pumps shuts off the test-section power.

Deionizer. A deionizer is in the system to maintain the water purity at a high value (resistivity approximately 2 megohm-cm). The deionizer loop draws about 1/2 gpm. It contains a bed of Amberlite MB-1 resin, 3 1/2 in. diam by 36 in. long, and two filters, a Neva Clog filter at the ion-bed inlet and a Micrometallic filter at its outlet. Automatic controls prevent the ion-exchanger from overheating during operation.

Preheater. In order to vary the inlet temperature to the test section, the loop has an immersion-type preheater with a capacity of 20 kw. Half of the power is controlled by switches and half by a hand-operated variac. The preheater cannot be turned on unless there is flow in the test section. The loop over-all temperature is controlled by 70 kw of "cast-in-bronze" line heaters. These heaters operate through the same type of safety circuit as the immersion preheater.

Heat Exchanger. There is a water-to-water heat exchanger in the system to remove heat from the primary flow.

Instrumentation

Burnout Detection. Incipient burnout of the test specimen is prevented from proceeding to actual failure by the use of a burnout detector or by observation of a wall-temperature excursion with a thermocouple. The burnout detector consists of the following:

(*a*) A bridge-type circuit of which the test specimen forms two legs.

(*b*) An amplifier that magnifies the bridge unbalance caused by a portion of the tube overheating.

(*c*) A thyratron that receives the amplified unbalance signal and strikes when the signal reaches a predetermined magnitude, permitting a condenser to discharge.

(*d*) A high-speed breaker whose trip coil is actuated by the condenser discharge.

Such a detection system is based on using a test-specimen material that has a large and continuous temperature coefficient of resistivity. When the test material meets these specifications, the detector will interrupt the power very close to the physical burnout point. When the specimen is constructed of a material whose resistivity does not increase with increasing temperature at high-temperature levels (for example, Zircaloy-2), the detector does not function properly and burnout must be detected either through a sudden, continuous rise in the wall temperature or through physical rupture of the specimen.

Burnouts have been checked with and without a detector. In all cases, the detector fired at heat-flux values greater than 90 per cent of the physical burnout flux.

Flow Meter. Flow is measured by a Potter turbine-type flow meter. The instrument is calibrated to be accurate to within 1/2 per cent of the instantaneous flow reading.

Temperature Measurement. Temperatures are measured by either a Leeds and Northrup 24-point precision indicator or a 16-point recorder. Both instruments have 10 suppressed ranges and are accurate to approximately 0.03 mv or slightly greater than 1 F at 635 F. All thermocouples are chromel-alumel.

Current-Measurement Shunts. Current is measured by the voltage drop across two calibrated 6000-amp Westinghouse type-G shunts. The readings are taken on a 5-range Brown recorder.

Voltage Measurement. The voltage drop across the test section is recorded on a 2-range Brown recorder.

Discussion

L. Bernath.[7] The authors are to be commended for undertaking so challenging an investigation. Perhaps a less ambitious scope of work would have yielded more consistent results. Inspection of the first set of data presented in Table 2 shows a lack of process control; for example, at constant subcooling and channel geometry, systematic reduction in coolant flow does not appear to result in a systematic decrease in burnout flux. When the data of Table 2, for the 0.055-in. channel with subcooling of 3 ± 3 deg F, are plotted as in Fig. 3, the best fit is obtained with a straight line of slope very nearly 0.9 and there appears no justification for fitting a curve to the data. If, in Fig. 3, the data points

[7] 416 Garland Road, Wilmington 3, Del.

from the nickel tube are eliminated (indeed, they should not be included since the ratio of heated surface to coolant flow area differs widely from that of the other channels), all but three of the remaining points lie in a very narrow region of the graph. Only by lending great weight to these three points can a relationship between heat flux and mass velocity be found, and the scatter of data points permits only a straight-line relation.

Again, in Figs. 4, 5, and 6, it is clear that straight lines fit each group of data points with more precision than do the curves of the authors. In addition, one might level the criticism that the curves obscure such trends in the data as become obvious from the *slopes* of the lines; e.g., both Figs. 4 and 5 show a greater slope for the 60 F than for the 10 F subcooled data. A cross plot between the slopes of the heat flux versus mass-velocity lines and the degree of inlet subcooling might uncover a useful relationship between the two.

The writer would like to interject a point at this time concerning the method of data presentation illustrated by Figs. 4 and 5. One cannot hope to describe burnout conditions at the downstream end of the test section by using inlet conditions (e.g., subcooling) for test sections of different geometrical configuration as the parameter. Clarification of the physical relationships which exist at burnout can be obtained only by consideration of the conditions at the site of the burnout, since burnout conditions result from the transition from nucleate to film boiling in a specific and limited region of the apparatus.

The data presented in Tables 5 and 6 purport to show the "possible" effects of inlet subcooling and L/D, respectively, on the burnout heat flux. These data merely show that one cannot determine the effect of a minor variable when the primary variable (mass velocity) is not held constant for the runs to be compared. There cannot possibly be an effect of L/D on the local burnout condition; in fact, the data in Table 6 clearly prove the validity of the first law of thermodynamics. These data show that, at constant inlet temperature of the coolant, if the flow rate is varied in direct proportion with test-section length, the burnout heat flux remains constant.

The writer believes that the first conclusion stated by the authors is incorrect. As pointed out previously, the curve presented in Fig. 3 is not a valid representation of the data from a single test-section geometry. The research work of McAdams,[8] Gunther,[9] and others, has shown clearly that a constant relationship exists between coolant velocity and the burnout heat flux for local boiling conditions. On the basis of established data, it must be pointed out that the extrapolation of results as performed by the authors leads to an erroneous conclusion.

It is unfortunate that the data reported by the authors were not collected with more care for they then would have been a worthy addition to the growing pool of knowledge in this field. However, as presented in this paper, neither these data nor the correlation stemming from them can be used with confidence by the reactor designer.

Authors' Closure

The authors wish to thank Mr. Bernath for his interest in this paper.

The reference to a lack of consistent results is questionable in that no theoretical basis exists to establish a priori the relationship between burnout and mass velocity at the conditions in this investigation. A preference for a straight-line relationship between heat flux and mass velocity in Fig. 3 at the expense of discarding some rectangular channel data, and the elimination of all nickel-tube results is difficult to ascertain. The comment that the ratio of heated surface to coolant flow area differs widely (in the case of the round tube) from the other channels is obviously incorrect, since the equivalent diameters of the 0.187-in. diameter tube and 0.097-in. rectangular channel are almost identical. The authors do not wish to imply, however, that there is any particular significance in the ratio of heated surface to coolant flow area.

With regard to Figs. 4, 5, and 6, there is nothing to indicate that straight lines fit the data better than the curves. A straight-line relationship between heat flux and mass velocity is used reluctantly by the authors in Equation [1], but only as an approximation.

If the discusser is implying that burnout can be correlated on the basis of local fluid conditions which have been calculated assuming complete mixing, this is only conjecture and is not borne out by any experiments which have been run to determine the effects of upstream conditions. In order to have a useful design equation, the authors have made the assumption of complete mixing and have used point conditions in obtaining Equations [1] and [2]; however, the influence of upstream conditions has been fully recognized and the L/D range for which the equations are applicable has been explicitly noted. A more recent paper by Roarty, et al.,[10] gives modifications to the present paper to account for a greater spread in L/D.

The mass velocity was not maintained constant in Tables 5 and 6 because of the infallibility of the First Law; however, the comparison sets generally indicate that the set of conditions having the higher mass velocity had the lower burnout flux. There is no reason to believe burnout is inversely proportional to mass velocity; therefore, the "possible" effects of L/D and inlet subcooling might be much greater than are shown herein.

The discusser's exception to the first conclusion of this study is true only if certain data, considered valid by the authors, are eliminated as discussed previously. It should be noted that the work of McAdams and Gunther, which is referred to, was specifically for low pressures with considerable subcooling at the burnout point. In no way do these data contradict the authors' first conclusion concerning approximately zero subcooled data. The authors cannot comment on "others" mentioned by the discusser.

The authors prefer not to comment on the alleged lack of care taken in collecting the data. It is felt that the reader, familiar with high-pressure burnout testing, will be able to judge this matter for himself.

It is the discusser's prerogative not to recommend the use of the work presented for reactor design. The authors must conclude that the basis for this reservation is contained in the discussion. Unfortunately this discussion does not support the reservation but merely reflects a misunderstanding of certain parts of the paper which should be cleared up by the authors' closure. During the past several years since the data of this report were taken, considerable electrically heated and in-pile burnout data have been obtained by Bettis Laboratory. No cases have been found where the equations presented in the present report were not conservative. It is unfortunate that security classification prevents more detailed descriptions of the use of the equations presented, in reactor design.

[8] "Heat Transfer at High Rates to Water With Surface Boiling," by W. H. McAdams, et al., *Industrial and Engineering Chemistry*, vol. 41, 1949, pp. 1945–1953.

[9] "Photographic Study of Surface-Boiling Heat Transfer to Water With Forced Convection," by F. C. Gunther, Trans. ASME, vol. 73, 1951, pp. 115 to 123.

[10] "Thermal Design Criteria for High Pressure Water-Cooled Reactors," by J. D. Roarty, W. M. Jacobi, K. M. Treadwell, N. C. Sher, and J. E. Zerbe, presented at ANS Meeting, Pittsburgh, Pa., June, 1957.

Properties of Friction Materials I—Experiments on Variables Affecting Noise

By P. R. BASFORD[1] and S. B. TWISS,[2] DETROIT, MICH.

Measurements of the friction of various brake linings against polished iron have been made at speeds so low that surface temperature (known to be important at higher speeds) could be neglected. Samples had to be run in thoroughly immediately before testing to secure reproducible results. The coefficient of dynamic friction was lower than the static coefficient at the lowest speeds, increased markedly, and approached a constant value at the highest speeds studied (12.5 ipm). Other apparatus was used to extend the range to 800 fpm. When the surface temperature was held constant, the friction coefficient passed through a broad maximum and thereafter decreased slightly as the speed was increased. The transition from smooth sliding to stick-slip friction was studied as a function of speed and load. A critical speed was found, above which only smooth sliding was possible, regardless of load. The time-force traces obtained during stick-slip motion supply information about the static coefficient, and also the elastic properties of the brake lining.

Introduction

The friction of metals (1, 2),[3] alloys (3), and some inorganic crystals (4, 5) has been studied extensively. The concept of the formation and shearing of welded areas (sometimes modified by plastic flow and strain-hardening) accounts very satisfactorily for the observed effects. However, the number of substances measured remains small. One class of materials about which little is known, in spite of their industrial importance, is brake linings. It is the purpose of this paper to develop experimental methods for the study of their frictional properties. A companion paper[4] examines the theory involved, with particular reference to their tendency to vibration in use.

The literature in the field, although extensive, offers so little help in this direction that it need not be examined in detail. Many inconsistencies are on record, both in the results and interpretation of friction tests; it appears that the results depend almost as much on the test method as on the material being tested. This is not surprising in view of the experimental difficulties. Failure to recognize the importance of the temperature at the rubbing surface and its dependence on the rate of heat generation seem to be responsible for much of the confusion. Under the proper conditions (to be discussed later) the surface temperature can be measured or calculated, but it cannot be estimated even approximately by means of a thermocouple located any distance from the surface. Mechanical factors, hard to analyze but sometimes important, also may obscure experimental results. Finally, the properties of brake linings are determined in part by the thermal history of the sample.

Apparatus and Methods

Three samples of brake linings used, A, C, and D, were commercial products, representative of original equipment linings. Seven others, B, and E through J, were formulated in this laboratory. The latter were so chosen that the effect of composition and processing conditions (closely guarded secrets for commercial materials) could be studied systematically. Table 1 gives some of the physical properties of these linings as well as the type of binder, which appears to be the most important composition variable affecting noise.

The apparatus shown in Fig. 1 was found to give reliable results

Table 1 Physical properties of brake linings

Designation	Binder type[a]	Density, g/cm³	Porosity, per cent	Hardness, Rockwell M	E_s, psi
A	R	2.24	5.2	92.6	1024
B	R-R	1.62	10.1	11.4	7361
C	Rb	1.78	15.8	16.5	5842
D	R	1.56	3.6	15.2	8935
E	R-R	1.54	17.0	−102.5	2108
F	R-R	1.54	16.9	−10.0	3406
G	R-R	1.60	10.8	20.0	6373
H	R-R	1.62	12.8	−55.3	7116
I	R-R	1.50	20.7	53.0	7258
J	R-R	...	...	...	9844

[a] R = Resin.
R-R = Resin-rubber.
Rb = Rubber.

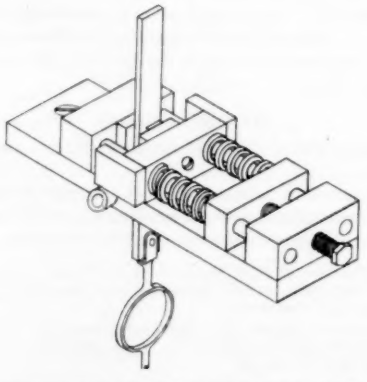


Fig. 1 Apparatus for friction measurements at low speeds

[1] Project Engineer, Engineering Division, Chrysler Corporation.
[2] Assistant Chief Engineer, Chemical Research, Engineering Division, Chrysler Corporation.
[3] Numbers in parentheses refer to the Bibliography at the end of the paper.
[4] "Properties of Friction Materials, II—Theory of Vibration in Brakes," by P. R. Basford and S. B. Twiss, published in this issue, pp. 407–410.

Contributed by the Lubrication Division and presented at a joint session of the Lubrication and Heat Transfer Divisions at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 8, 1956. Paper No. 57—SA-96.

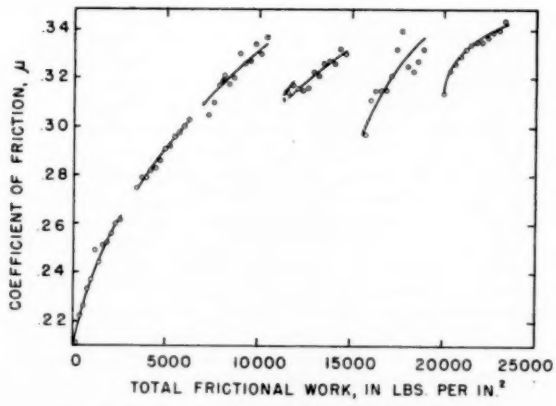


Fig. 2 Increase of friction during run-in of Sample A

at the very low speeds for which it was designed. Essentially, it consists of a specially ground (5–10 μ-in. rms) iron or steel bar clamped between two samples of brake lining and pulled through at a predetermined rate. The bar was anchored through a strain ring to the fixed head of a Tinius Olsen tensile machine, while the sample fixture was mounted on the movable crosshead. Speeds between 0.025 and 12.4 ipm could be selected and maintained by means of a balanced thyratron circuit which controlled the driving motor. The normal (load) force on the sample was supplied by calibrated springs compressed by a screw-operated block. The length of the springs under compression was used to measure the normal force. The output of the strain gage was fed into a Brown recording potentiometer, which supplied a continuous record of the frictional force during a test. The metal surfaces were not cleaned with solvents, nor was contact permitted with anything except the lining samples. They were stored in a desiccator when not in use. Under these conditions, the metal surfaces may be considered clean, except for the thin oxide films which always form in contact with air.

Friction was always low and erratic for freshly ground bars and samples. Fifty to one hundred traverses of the surface resulted in a gradual increase of friction, but not to a constant value. Additional run-in on succeeding days increased friction further until, at the end of five or six days, a substantially constant coefficient of friction was attained. This behavior, which is observed only for rubber and resin-bonded materials, depends on the gradual transfer of a thin film of binder to the surface of the metal. Fig. 2 shows the increase of the coefficient of friction as Sample A was run in. The abscissa is the total frictional work done up to the time of the measurement. One material was encountered (Sample D) which could not be stabilized even with twice the amount of run-in found to be adequate for other samples.

When the sample was run in properly, friction measurements were started, using a series of increasing loads at a constant rubbing speed. Typical results, for Sample A at 12.4 ipm, are shown in Fig. 3. The circles and triangles refer to two independent, consecutive sets of measurements. The coefficient of friction is equal to one half the slope of the line, as determined by least squares. The scattering of points was always small. However, the method is less accurate than this would indicate, since duplicate measurements differed, on the average, by 0.01 to 0.02 unless they were done consecutively. Inability to control moisture adsorption and formation of surface oxide on the metal is probably responsible for this variation. Similar measurements were made at various other speeds for each sample.

Two distinct kinds of motion were observed: At loads less than a critical load (which depended on the rubbing speed) the samples slid smoothly, but when the critical load was exceeded, a sharp transition to stick-slip motion took place. The relatively slow response of the Brown recorder makes it unsuitable for study of the high-frequency, low-amplitude vibrations over the transition region. A Brush recorder capable of responding to 120 cps was used to determine the critical load for various rubbing speeds. A high-gain preamplifier was necessary. The load, originally greater than the critical load, was decreased gradually, the frequency and amplitude being noted after each decrement. A plot of amplitude versus load could be extrapolated to zero amplitude to give the critical load.

The actual transition region could be observed only with an oscilloscope, also with high preamplification. Two characteristic wave forms were noted; namely, a saw-tooth wave during stick-slip motion, and a highly irregular form due to the random unsynchronized force increments during smooth sliding. The relation of these experiments to the general problem of noise and vibration are discussed in the companion paper[4] in some detail.

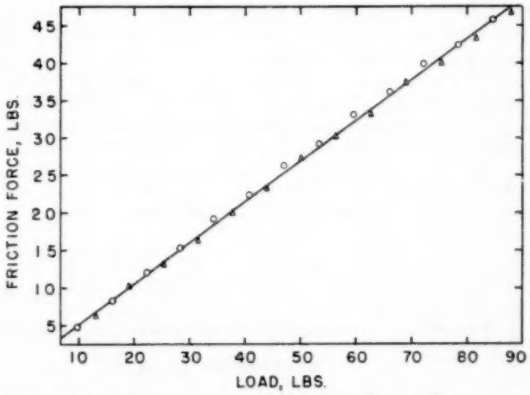


Fig. 3 Typical measurement of dynamic coefficient at low speeds, Sample A at 12.4 ipm

At loads well above the critical, the Brown recorder could reproduce the low-frequency saw-tooth wave form adequately. The results were used to determine the static coefficient of friction. The maximum tangential force recorded is equal to load $\times$ static coefficient. The slope of a plot of maximum force versus load is equal to twice the static coefficient. A typical example, for Sample A, is shown in Fig. 4. Data for the dynamic coefficient at 0.031 ipm are included for comparison. The static coefficient so measured showed a tendency to increase slightly as the sticking time increased.

In addition, the force traces during stick-slip motion provide information about the resistance for the lining to deformation by shear forces, a property closely related to its tendency to vibration in use. When the traces are replotted with the distance between slips as the abscissa, it was found that the slope, i.e.

$$\left(\frac{df}{dl} \quad \text{or} \quad \frac{d(\text{tangential force})}{d(\text{displacement})}\right)$$

was characteristic of the sample and did not change with speed or load. The elastic constant in shear E_s is determined from

$$E_s = \frac{t}{A}\left(\frac{df}{dl}\right) \qquad [1]$$

where t = sample thickness, A = contact area, and df/dl is the

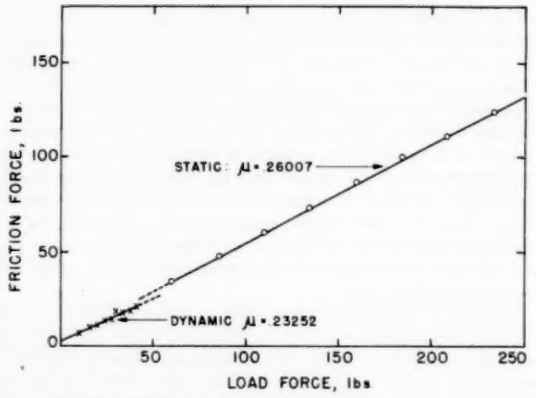


Fig. 4 Comparison of static and dynamic (0.031 ipm) coefficients, Sample A

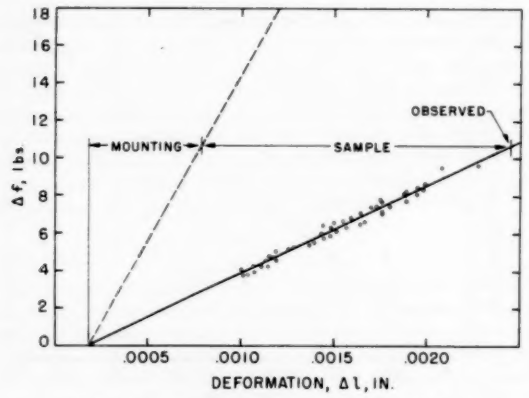


Fig. 5 Deformation under stress, Sample A

rate of increase of the tangential force as the displacement l increases during the stick part of the cycle.

The measured displacement consists of two parts: (*a*) Distortion of the sample under shear; and (*b*) distortion of the strain ring and mounting under tension. The latter was measured separately, and the results used as a correction factor. When Δf is plotted against Δl, straight lines such as those shown in Fig. 5 always resulted. From them E_s is determined

$$E_s = \frac{t}{A\left[\left(\frac{df}{dl}\right)_{obs} - \left(\frac{df}{dl}\right)_{corr}\right]} \quad \ldots\ldots\ldots\ldots [2]$$

Considerable variation in E_s is observed for the different linings investigated as shown in Table 1.

The apparatus described in the foregoing cannot be used to study the effect of temperature or higher rubbing speeds. A laboratory friction machine shown schematically in Fig. 6 is quite suitable for such experiments because of its extreme flexibility (6). It consists of a horizontal cast-iron disk specially surfaced to a 3–6 μ-in. rms finish and rotated at a constant predetermined speed by a geared-down electric motor. Rubbing speeds between 25 and 1600 fpm are attainable. A sample of brake lining $^1/_2 \times 1$ in. rubs on the upper surface. It is mounted on a horizontal arm pivoted at one end to allow both horizontal and vertical rotation.

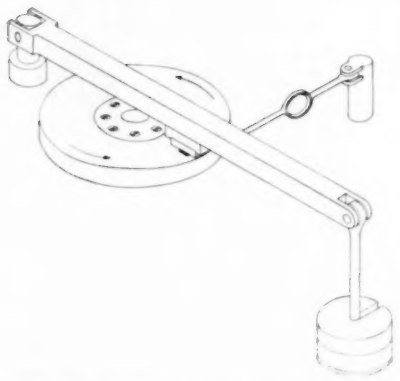


Fig. 6 Schematic diagram of laboratory friction machine

Horizontal movement is restrained by a strain ring which actuates a Brown recording potentiometer. The load force is supplied by weights suspended from the free end of the arm. An electrical heater and a thermocouple are imbedded in the disk. The temperature can be controlled either by a Leeds and Northrup potentiometer operating from the thermocouple, or by manual adjustment of a powerstat in the heater circuit.

To ensure mating surfaces and a stabilized coefficient of friction, all samples were run in 15 hr at 350 fpm and 111 psi load.

It is known that the coefficient of friction of brake linings depends markedly on the temperature at the rubbing surface. Ideally, therefore, high-speed friction should have been measured at room temperature, as the low-speed friction was. Unfortunately, so much frictional heat is generated that this is impossible without artificial cooling. The course adopted was to determine the effect of speed at a series of temperatures from 185 to 435 F, using the utmost care to hold the temperature constant to within 5 deg for each set.

This could be done in two ways. It was found that if frictional heat was generated at a constant rate $\dot{q}$ the surface temperature remained substantially constant regardless of speed. Consequently, it is only necessary to compensate any increase of speed by that reduction of load which will restore $\dot{q}$ (i.e., $L\mu v$, where L = load, μ = coefficient of friction, and v = linear speed) to its former value. Minor adjustment of either v or the rate of electrical heating will then bring the temperature to its preassigned value.

Before using this method, it is necessary to be sure that the coefficient of friction is independent of load, as required by Amontons' law. To establish this point, several sets of measurements were made at constant speed, and at temperatures high enough so an increase of load could be compensated by a decrease in the electric heating. In every case the plot of friction force versus load was strictly linear, as expected.

It was found that the disk thermocouple, which is located $^1/_4$ in. below the friction surface, was not suitable for measuring surface temperatures when the bulk of the heating came from friction rather than the electric heater. Two very small iron-constantan thermocouples were mounted in the brake-lining samples, one about 0.010 in. from the surface, the other close to the mounting block. Normally, the temperature gradient within the sample was linear, and the surface temperature could be obtained by simple extrapolation. Rapidly changing surface temperatures could be estimated by an approximate solution of the heat-flow equation, using the time derivatives of the two temperatures and assuming that the second derivative of the temperature with respect to distance is linear. No experiments of this kind are included here.

Results and Discussion

It was noted previously that friction increased gradually as run-in proceeded. The reason for this is brought out by a series of experiments on static friction. The coefficient of friction of Lining B was 0.16 originally; thorough run-in brought this up to 0.41. Cleaning the metal surface with toluene reduced the coefficient to its original figure. Neither lapse of time, heating, nor polishing with very fine abrasive (any of which should remove a residual film of solvent) served to restore the friction to its former level. Significantly, it could be restored partially by treating the metal surface with a toluene extract of the same kind of lining, followed by heating to drive off the solvent. It appears from this that the increase of friction during run-in is associated with transfer of a film of organic binder to the metal surface. Such films can be observed in a microscope under grazing illumination as hazy, structureless brown layers.

The original friction is exclusively mechanical (i.e., abrasive) in nature; presumably it depends on the properties of the asbestos and the roughness of the metal surface. In the presence of a transferred film, the abrasive friction can be supplemented by formation and shearing of joints between the binder and the transferred film, an effect governed solely by the properties of the binder. To judge from the measured contribution to the coefficient—0.16 for abrasive friction, 0.25 for bond-shearing friction—the latter is more important at room temperature by a factor of approximately 3:2.

Present experimental evidence does not warrant a detailed discussion of the two kinds of friction. Qualitatively, however, the data conform in several respects to what would be expected on the basis of this distinction, as will be pointed out.

The results obtained at low speeds under conditions of smooth sliding are shown in Fig. 7. The static coefficients are included for comparison. This diagram shows two features: (*a*) The coefficient of friction is low at the lowest rubbing speeds, but tends to approach a higher, constant value as the speed is increased. This increase seems to be a general effect. (*b*) The static coefficient is significantly higher than the apparent limit of the dynamic coefficient as the rubbing speed approaches zero. The known properties of the binders would lead one to expect exactly what is observed. The binders are completely amorphous or glass-like; they will, therefore, when highly stressed as at the areas of contact, behave not as crystalline solids but as very viscous liquids. As such they will flow under the action of normal (load) forces so as to relieve the stress, the result being that the bonded area is increased. The effectiveness of this process will depend on the time allowed for flow; i.e., the time the areas remain in contact. Contact time, flow, and bonded area being greatest under static conditions, the static coefficient will be high in relation to the dynamic, as shown in Fig. 7. A second and presumably more important effect, synchronization of joint rupture, is discussed in the companion paper[4] in connection with the theory of noise.

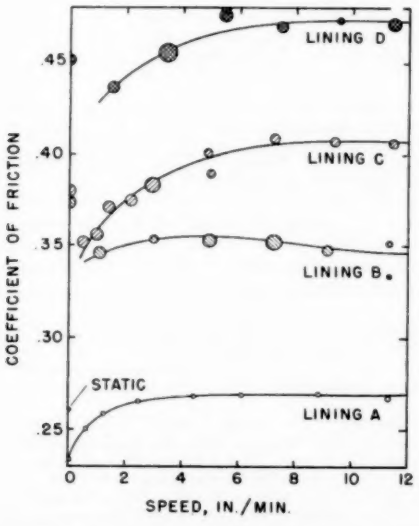


Fig. 7 Coefficients of friction at low speeds, room temperature

Flow is to be expected under tangential as well as normal stresses. This was demonstrated experimentally. At a speed of 0.025 ipm and a load of 220 psi, stick-slip friction gave rise to the saw-tooth recorder traces shown in Fig. 8 for Sample C. If the machine was stopped while the sample was sticking, the tangential force decreased as shown. This behavior is not characteristic of the machine; when a spring was substituted for the sample assembly the force remained constant indefinitely. The rate of shear could not be measured directly, but is estimated to be about 2×10^{-4} ipm. It cannot be assumed that the binder will flow tangentially at all rates of shear; more commonly, such materials behave like viscous liquids at low shear rates, and like brittle solids at high shear rates. The energy necessary to rupture a joint, and consequently the friction force, is presumably higher in the latter case. The observed increase in friction with rubbing speed is believed to be associated with the gradual transition from viscous liquid behavior to brittle solid behavior which takes place at the contact areas as the rate of shear increases.

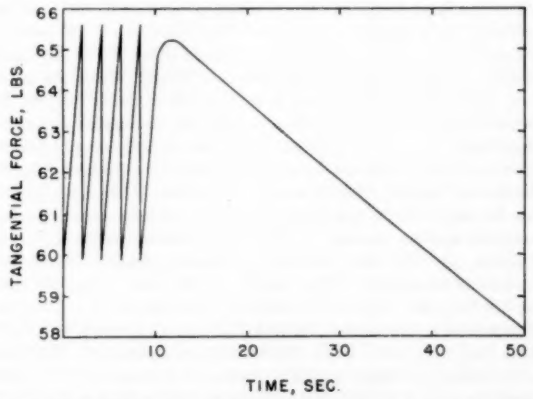


Fig. 8 Stick-slip friction and flow under tangential stress, Sample C

The data obtained on the friction machine show the effect of rubbing speed over a much wider range. Comparison with the low-speed results is hampered somewhat by inability to maintain low surface temperatures at the high rates of heat generation unavoidable in high-speed experiments. The over-all pattern is unmistakable, however; this is shown in Fig. 9 for several linings. The surface temperature was held constant at 300 ± 5 deg F by varying the input to the disk heater, as discussed earlier. A marked increase in friction at the lowest speeds is followed by a long region of slightly increasing friction, a broad maximum, and at the highest speeds, a slightly decreasing coefficient. The decrease in friction is small above 200 fpm, and might almost be ignored except for its bearing on the tendency of a lining to be noisy in use.

It was noted previously that a sharp transition between smooth sliding and stick-slip friction always was observed. A typical ex-

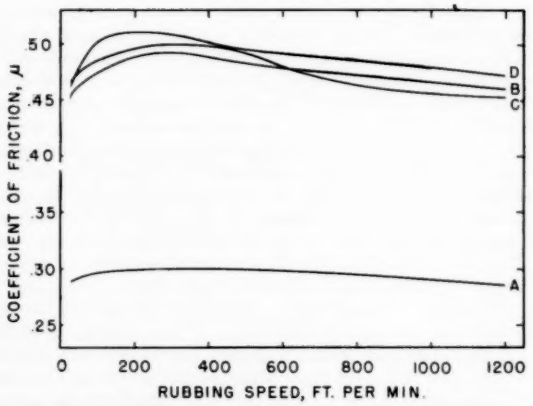


Fig. 9 Coefficients of friction at high speed, temperature constant at 300 F ± 5 deg F

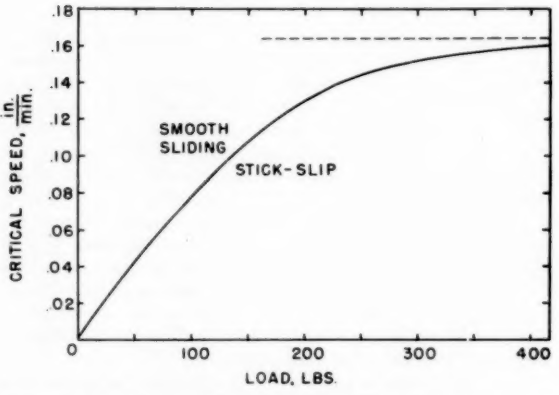


Fig. 10 Transition from stick-slip to smooth sliding, relation between load and critical velocity, Sample A

ample, for Sample A, is shown in Fig. 10. Two points are evident here; there is a critical velocity above which no load, however great, can induce stick-slip motion, and (by reference to Fig. 7) that the dynamic coefficient is not equal to the static coefficient at the critical speed, 0.164 ipm, and does not become equal to it until the materially higher speed of 1.546 ipm is reached. The usual requirement for stick-slip friction, that the static coefficient exceed the dynamic coefficient, is therefore a necessary but not a sufficient condition. The fact seems to be that stick-slip motion is not determined solely by the frictional properties of the rubbing surfaces. A progressive increase of the stiffness or the mass of the mounting system results in lower amplitudes of vibration and smaller cyclic variations in the frictional force. When the latter falls within the range of random-force fluctuations introduced by the statistical nature of friction, the motion no longer can be distinguished from smooth sliding. A related effect depends on the time τ necessary to break the group of bonds responsible for static friction during the sticking part of the cycle. This is small, but can be observed on the Brush recorder. When taken into account in a time average, it will lower the effective static coefficient, or raise the dynamic coefficient, depending on the method of averaging. The equality of the two coefficients thus modified appears to limit the region of stick-slip motion. In terms of measurable quantities, the ratio of τ (which changes very little with load and speed) to T, the duration of a complete stick-slip cycle (which changes radically with speed, load, rigidity, and mass) must be less than some critical value for stick-slip friction to occur.

Whatever the explanation may be, these experiments show that stick-slip friction in the usual meaning of the term cannot be responsible for vibration during high-speed operation of brakes.

In Table 1 are given the results of the measurements of the shear elastic constant. These were calculated by use of Equation [2] and force versus displacement data, as shown in Fig. 5 for Sample B. There is some uncertainty in the placing of the line which determines the correction for distortion of the mounting, so that the E_s-values cannot be accepted as final. They are, however, comparable among themselves, and quite adequate for the qualitative treatment intended here.

The results discussed in the foregoing supply much of the information needed to develop a comprehensive theory of noise and vibration in brakes. This is discussed in the companion paper.[4]

Bibliography

1 "The Friction and Lubrication of Solids," by F. P. Bowden and D. Tabor, Oxford University Press, London, England, 1950.

2 "Dry Metallic Friction as a Function of Temperature Between 4.2° K. and 600° K.," by I. Simon, H. O. McMahon, and R. J. Bowen, *Journal of Applied Physics*, vol. 22, 1951, pp. 117–184.

3 "Solid Solubility Effect of Metallic Surface Friction," by K. Umeda and Y. Nakano, *Journal of the Faculty of Science*, Hokkaido University, series II, vol. 4, no. 1, 1951, pp. 70–86.

4 "Friction of Diamond, Graphite, and Carbon: The Influence of Adsorbed Films," by F. P. Bowden, J. E. Young, and G. Rowe, Proceedings of the Royal Society of London, series A, vol. 212, 1952, pp. 485–488.

5 "The Friction of Non-Metallic Solids," by F. P. Bowden, *Journal of the Institute of Petroleum*, vol. 40, 1954, pp. 89–101.

6 "New Apparatus for Friction Measurement," by P. J. Willson, S. B. Twiss, and D. M. Teague, *ISA Journal* (Instrument Society of America), vol. 3, 1956, pp. 224–228.

Properties of Friction Materials II—Theory of Vibration in Brakes

By P. R. BASFORD[1] and S. B. TWISS,[2] DETROIT, MICH.

A theory of vibration in brakes is developed, based on the statistical nature of friction. The conditions under which an incipient vibration can develop are shown to be (*i*), $4\,bm > a^2L^2$ where b = elastic constant of the lining in shear, m = mass per unit area, a = change in coefficient of friction with speed, and L = load force per unit area; and also, (*ii*), $a < 0$.

Whether noise will result from the vibration depends on how close the natural frequency of the lining, $(4\,bm - a^2L^2)^{1/2}/(4\pi m)$, is to a frequency of the drum which can be excited by resonance.

Observations on the relative noisiness of four kinds of lining were correlated with measurements of a and b. When the linings are ranked in order of increasing tendency to noise as predicted by the theory, it is found that the order is the same as that observed in brake tests on road cars.

Introduction

THE importance of noise in brakes is reflected in the large body of published work on the subject. It is not necessary to review this in detail, since, with a few exceptions noted later, the emphasis is on practical corrective measures to the exclusion of theory. Two approaches have been used: (*a*) Analysis of mechanical factors—rigidity, natural frequencies, and self-damping of the system, completeness of contact between lining and drum, and so on; and (*b*) study of the characteristics of the lining which are associated with vibration.

The first has not led to conclusive results; no consistent relation between vibration and any mechanical property has ever been demonstrated. One significant fact is on record; the natural frequencies of brake systems are so numerous and closely spaced that almost any inciting vibration will be amplified by resonance.

The group of papers on lining properties indicates that vibration is related in some way to the stick-slip phenomenon, and also to a coefficient of friction which falls as the rubbing speed increases. The nature of this relation was first clarified by Dudley and Swift (1),[3] who showed that vibration was possible without actual sticking. They also showed that energy in excess of a certain threshold value must be supplied before vibration by resonance, or otherwise can be built up to audible levels. This was confirmed experimentally by Sinclair (2), who observed wave forms and coefficients of friction at low speeds. The nature of the vibration is determined jointly by the frictional properties of the lining and by the elastic properties of the mounting, which makes interpretation difficult. Lining inhomogeneity and high coefficients of friction also have been held responsible for vibration, although the evidence seems inconclusive.

To clear up this confusing situation, it is necessary to develop and test a comprehensive theory of vibration in brakes. Enough is now known of friction phenomena in general to make this possible.

Development of Theory

A valid theory of vibration in brakes must account for a number of observations not obviously related:

1 Vibration occurs erratically, appearing and disappearing for no apparent reason.

2 In general, the frequency does not depend on speed or load.

3 "Ovaling" vibration of the drum, with tangential movement at the nodes and radial movement at the antinodes is responsible for noise. The number and spacing of nodes is variable.

4 When tested on the same brake, a series of linings can be ranked qualitatively according to their tendency to noise, ranging from almost uniformly quiet to almost uniformly noisy.

5 Linings with substantially equal coefficients of friction may be very different in their tendency to noise.

6 A single lining may be quiet on one brake and noisy on a brake of different design.

7 In borderline cases, noise can be suppressed by springs wrapped around the drums or other damping devices.

Certain of these observations—items 2, 3, 6, and 7—depend on mechanical factors without reference to any properties of the lining. In principle, it should be possible to predict the frequency and wave form of vibrations from the shape, mass, and elastic properties of the system, and, by considering the degree of self-damping to estimate the probability of noise. In practice, it is virtually impossible because of the complexity of the problem; results either must lack generality, be based on questionable assumptions, or be too cumbersome to be useful.

This is less serious than it appears. It is certain from observation 7 that mechanical factors are decisive only in borderline cases. Observations 4 and 5 indicate that, in general, the probability of noise is determined jointly by lining properties and mechanical factors, with the former much more important. Observation 1, on the random occurrence of noise, suggests that vibration is initiated in the lining; otherwise it is difficult to see why identical mechanical factors operating during each of a series of comparable stops should not produce identical results; i.e., all stops noisy or all stops quiet.

Two points must be established before the foregoing tentative conclusions can be accepted: (*a*) The erratic and unpredictable incidence of vibration must be shown to be deductible from known properties of the linings; and (*b*) the mechanism by which vibration is initiated and built up to audible levels must be demonstrated and shown to furnish a valid basis for correlation between measured properties of the linings and their observed tendency to noise.

Before discussing the first point, it will be well to specify as closely as possible exactly what is occurring randomly. An incipient vibration is characterized by an excess of potential energy somewhere in the system which is to be translated into kinetic

---

[1] Project Engineer, Engineering Division, Chrysler Corporation.

[2] Assistant Chief Engineer, Chemical Research, Engineering Division, Chrysler Corporation.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Lubrication Division and presented at a joint session of the Lubrication and Heat Transfer Divisions at the Semi-Annual Meeting, San Francisco, Calif., June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 8, 1956. Paper No. 57—SA-97.

energy. Presumably, the excess energy is caused by a local and transient excess of frictional force, which has displaced a portion of the lining (area and location unspecified) from its equilibrium position against the restoring forces of elasticity. The randomness with which vibrations are initiated is therefore a consequence of the random-force fluctuations characteristic of the friction process.

It is noted in a companion paper[4] that smooth sliding was never observed when the measuring instrument was fast enough and sensitive enough to give true readings of the instantaneous friction force. When an oscilloscope was used, the trace showed extremely rapid, patternless fluctuations. Calibration with known forces indicated that the average band width, i.e., the average deviation of the friction force, was about $\pm 1^1/_2$ lb. This increased slightly with load, but was not affected by speed. Since the traces were not photographed, more exact estimates cannot be made.

Rabinowicz and co-workers (3, 4) using a recording system better adapted to quantitative work, have observed similar fluctuations with metallic friction. They have shown how they vary with experimental conditions and even, by an ingenious statistical method, have deduced some of the properties of the individual welds responsible for the fluctuations (5). They report standard (i.e., root-mean-square) deviations 2 to 5 per cent of the friction force. This is the same order of magnitude as that for the brake linings. Therefore it must be accepted as an established fact that friction forces, by the nature of the friction process itself, are subject to random fluctuations.

Applying this concept to the lining in operation, we see that a multitude of minute concentrations of potential energy must be formed continually. This poses two further questions: (*a*) Why are the great majority of these incapable of developing into vibrations? (*b*) How does the occasional energy concentration which does develop differ from the others? To answer these, we need only note that the friction forces are subject to random fluctuations with respect to position as well as time; that is, an element of area capable of passing into vibration by reason of its displacement may experience a sudden decrease of frictional force (favorable to the transition) or a gradual decrease (unfavorable). Moreover, if it does pass into vibration, it cannot execute independent vibrations; it is coupled, in the strictest sense, with adjacent elements of area whose amplitudes and phases, also being distributed at random, are more likely to damp out the incipient vibration than to reinforce it.

To overcome the inherent self-damping properties of the lining, a displacement must involve enough adjacent areas to ensure that its frictional properties are substantially those of the average of the lining, with random fluctuations canceling out each other. Under these conditions, the subsequent history of the vibration—either self-reinforcement or self-damping—is rigorously determined by the properties of the lining, as will be shown in the next section. For the present, we note only that simultaneous displacement of such an area is extremely improbable. Nevertheless, it is a property of random distributions that even such an improbable event is virtually certain to occur if enough time is given, and to recur at unpredictable intervals thereafter.

The erratic incidence of noise in brakes is therefore to be considered a necessary consequence of the accepted mechanism of friction—formation and shearing of discrete bonds with the attendant fluctuation in frictional force.

In order to determine what properties of brake linings are responsible for their notable differences with respect to noise, we consider a small area of lining which, having been displaced, is now vibrating in the x (tangential) direction. When no forces act on it, its equilibrium position is at $x = 0$. When it is at rest, the drum rubs against it with a speed V_0 in the positive x-direction; if it is moving at a speed (dx/dt), the relative speed falls to $(V_0 - dx/dt)$. The vibrating mass is m (per unit area) and the force holding it against the drum is L, also per unit area. Two forces act on it:

1 An elastic force proportional to the displacement

$$\left.\begin{aligned} f_e &= -E_s/t \cdot x \\ \text{or for convenience} \quad f_e &= -bx \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots [1]$$

where E_s is the shear elastic constant and t is the thickness of the lining. The minus sign indicates that this force tends to decrease motion in the positive x-direction.

2 A frictional force whose magnitude depends on the relative speed of lining and drum

$$f_f = L\phi(V_0 - dx/dt) \ldots\ldots\ldots\ldots\ldots [2]$$

L being the load per unit area. The sign here is positive because the force tends to increase motion in the positive x-direction. The function $\phi(V_0 - dx/dt)$, i.e., μ, is known to pass through a maximum and to decrease over much of the speed range. As an approximation we may use for $\phi(V_0 - dx/dt)$

$$\mu = \mu_0 + a(V_0 - dx/dt) \ldots\ldots\ldots\ldots\ldots [3]$$

i.e., μ decreases linearly with the relative velocity. This will be quite satisfactory when a narrow band of velocities is involved, as in the present case of an incipient vibration, but will cease to apply when the band widens with increasing amplitude. The friction force is then

$$f_f = L(\mu_0 + aV_0) - aL(dx/dt)$$

The equation of motion of the vibrating lining is

$$\left.\begin{aligned} m(d^2x/dt^2) &= L(\mu_0 + aV_0) - aL(dx/dt) - bx \\ \text{or} \qquad & \\ (d^2x/dt^2) + aL(dx/dt)/m + bx/m &= L(\mu_0 + aV_0)/m \end{aligned}\right\} \quad \ldots [4]$$

This equation can be integrated by any one of several methods; the solution is

$$x = A \cdot \exp(-aLt/2m) \sin [(4bm - a^2L^2)^{1/2}t/2m + B] \ldots [5]$$

The arbitrary constants A and B are determined by the amplitude and phase, respectively, when $t = 0$; for our purpose there is no need to evaluate them.

Before attempting to use Equation [5], it will be well to consider a condition under which it could not be valid. If the maximum velocity attained during vibration becomes equal to or greater than V_0, the friction force is reversed in direction, the stick-slip phenomenon may be induced, and the resulting discontinuity makes any analytic solution impossible. Very likely Equation [5] cannot describe the later stages of build-up for this reason, but for the present purpose—studying the incipient vibration during the short but all-important period when its subsequent history (self-damping or self-reinforcement) is determined—the limitation does not apply. It is only necessary to show that, initially, a reasonable choice of parameters leads to a maximum vibrational velocity small in comparison with the V_0's to be expected during operation of a brake.

By ordinary methods Equation [5] can be put into the form

$$(dx/dt)_{max} = 2\pi\omega A \exp(-aL/8\omega m) \ldots\ldots\ldots [6]$$

If the initial displacement A is assigned the rather high value

[4] "Properties of Friction Materials, I—Experiments on Variables Affecting Noise," by P. R. Basford and S. B. Twiss, published in this issue, pp. 402–406.

0.005 in., ω set equal to 2300 cycles per sec (a commonly observed frequency), and the damping term exp $(-aL/8\omega m)$ neglected, i.e., assumed to be 1, the maximum vibration velocity is no more than 72.3 ips. At all car speeds above 9 mph, V_0 is greater than this. Therefore the limitation can be ignored.

The properties of a lining which determine its tendency to noise are apparent from the form of Equation [5]:

1 An initial displacement can develop into vibration only if $4bm > a^2L^2$; otherwise it will decrease exponentially with time.

2 If $a < 0$, the vibration will be self-reinforcing, the associated energy eventually will exceed the threshold level, and noise will result. If, on the other hand, $a > 0$, the exponential term leads to self-damping, since L, m, and t are all necessarily positive.

Recognition of the importance of the shear elastic constant of linings opens up the possibility of suppressing noise at the source by modifying this property.

To complete the analysis, we must bridge the gap between Equation [5], involving only the lining, and the final vibration, which involves the drum as well. It is obvious that during the build-up period whatever forces operate on the lining must operate equally (in the opposite direction, of course) on the drum. The limited area over which this energy transfer from lining to drum takes place probably becomes a node for the final vibration. The possibility of a transfer area developing anywhere on the lining accounts for the random positioning of nodes which has been observed.

During the build-up period, the drum and lining execute coupled vibrations; i.e., the natural vibration of either is disturbed by the tendency of the other to do something quite different. Specifically, if the inciting vibration is close to a natural frequency of the drum, build-up will be faster. Conversely, if the drum is equipped with effective damping devices, build-up will be slower and in a borderline case may be inhibited altogether. Formally, Equation [5] can be modified to take coupling into account by inclusion of a term D, characterizing the behavior of the drum at the inciting frequency

$$x = A_e - \left(\frac{aL}{2m} + D\right) t \sin\left[\frac{(4bm - a^2L^2)^{1/2}}{2m} t + B\right] \ldots [7]$$

D is defined here simply as the algebraic sum, normally small, of damping terms (positive) and resonance terms (negative), to which we will refer later.

The foregoing analysis of vibration in brakes leads to the following conclusions:

1 The erratic and unpredictable occurrence of noise is a direct consequence of the random fluctuations of friction force inherent in bond-shearing friction.

2 Self-reinforcement of incipient vibrations is possible only when the coefficient of friction decreases with speed.

3 Vibration of any kind is possible only when $4bm$ is greater than a^2L^2.

4 Both a and b vary widely from lining to lining; hence the wide differences observed in their tendencies to noise.

5 The randomly positioned nodes of the fully developed vibration probably mark the places where vibration was initiated.

6 The effectiveness of damping devices in borderline cases depends on the coupling of vibrations in lining and drum during build-up.

Fig. 1 shows the joint effect of a and b (i.e., $d\mu/dV$ and E_s/t) in determining vibration. Fig. 2 shows the form of vibration to be expected over various regions of Fig. 1.

In the foregoing discussion we have attempted to determine the conditions under which noise is possible and something of the mechanism by which it is initiated and built up. The factors which make noise more probable in one lining than another must now be considered.

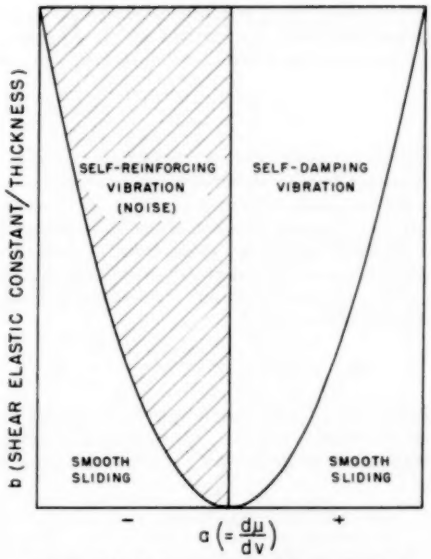


Fig. 1 Joint influence of a and b on vibration

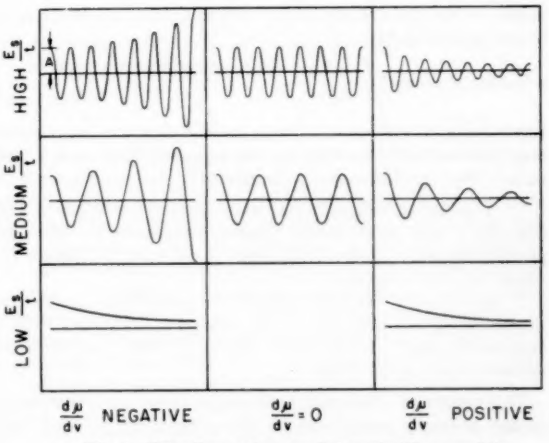


Fig. 2 Vibration pattern as determined by a and b

One factor is obviously the rate at which the vibration energy builds up, since a rapid increase leaves less time for the operation of the normal self-damping noted. From Equation [6], the rate can be shown to be

$$\left(\frac{dE}{dt}\right)_{t=0} = -2E_0 \frac{aL}{2m} \ldots\ldots\ldots\ldots\ldots [8]$$

where E_0 is the local and transient excess of potential energy which is to be built up; it is arbitrary except that it must be held constant to insure fair comparisons between different linings. Therefore the probability of noise will be proportional to $-aL/2m$, other things being equal.

The constant b does not appear in Equation [8]. Nevertheless, it plays an important part in determining the probability of noise. This hinges on the fact, noted previously, that the

coupled vibrations of lining and drum are amplified by resonance if the natural frequencies of the two are close together. The natural frequency of the lining vibration is, from Equation [6]

$$\omega = \frac{(4bm - a^2L^2)^{1/2}}{4\pi m} \quad \ldots\ldots\ldots\ldots\ldots [9]$$

That of the drum cannot be specified exactly. The large number of closely spaced frequencies which has been reported must refer to many different modes of vibration, only a few of which can be excited by the tangential vibration of the lining, and only one of which is ordinarily excited in a given drum. It is therefore permissible to use ω_{noise} as the parameter characterizing drum vibration, in which case the contribution of resonance to energy build-up is a function of the variable

$$\Delta\omega, \text{ i.e. } \left[\omega_{noise} - \frac{(4bm - a^2L^2)^{1/2}}{4\pi m}\right]$$

The nature of this function cannot be deduced for such a complex system as the brake assembly. It is reasonable to assume, however, that it resembles the normal (Gaussian) distribution curves

$$f(\Delta\omega) = \frac{1}{(2\pi)^{1/2}\sigma} e^{-(\Delta\omega)^2/2\sigma^2} \quad \ldots\ldots\ldots\ldots [10]$$

where σ is determined by the width of the frequency band which can induce resonance.

The discussion may be summarized as follows: (Probability of noise) = (constant) (rate factor) (resonance factor), which may be approximated by

$$p\,(\text{noise}) = (\text{const})\left(\frac{-aL}{2m}\right) e^{-[\omega_{noise} - (4bm - a^2L^2)^{1/2}]/(2\sigma^2)} \quad \ldots [11]$$

Fig. 3 is a plot of Equation [11] in arbitrary units to show qualitatively how the probability of noise depends jointly on a, b, and ω_{noise}. Two conclusions are to be drawn from this:

1 A group of linings having substantially the same a, i.e., $(d\mu)/(dV)$, may show widely different tendencies to noise, depending on their elastic properties; those with the highest E_s-values will be the noisiest.

2 If the group has substantially the same elastic properties, the probability of noise will be zero when $a = 0$ and also when a approaches some critical (more negative) value. Intermediate values of a will lead to noise.

This completes the development of the theory. It must be regarded as largely qualitative for the present, but as valid data accumulate, it can be made quantitative. In any case, it supplies a rational guide for further experiments.

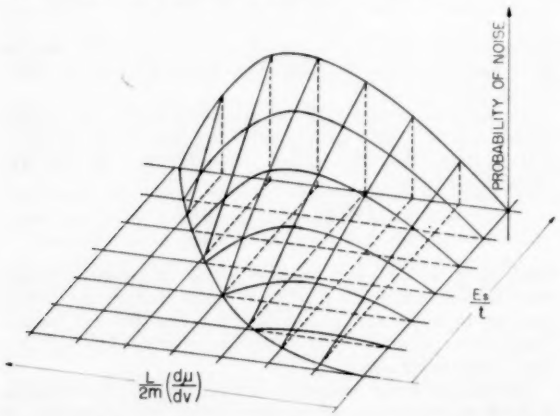


Fig. 3 Combined effect of a and b on probability of noise

Table 1 Results of tests in four linings

	$\frac{(d\mu)^*}{(dv)}$, $\frac{\text{sec}}{\text{cm}}$	E_s, psi
Lining A	−0.004675	1024
Lining B	−0.006835	5842
Lining C	−0.005906	7361
Lining D	−0.006288	8935

* Measured at 900 fpm and 150 C.

Comparison Between Theory and Experiment

Every step in the development of the theory was determined by the necessity of conforming to observed facts. Several points of correspondence have been noted already—the random occurrence of noise, the random spacing of nodes, the lack of correlation between noise and coefficient of friction, and so on.

Perhaps the most critical test of the theory is its ability to predict the tendency of a lining to be noisy in use on the basis of measurable physical properties. For this purpose, some recognized measure of noise (probability of occurrence, as distinct from intensity) is necessary. Unfortunately, the results of brake testing have never been recorded in this way, so quantitative predictions cannot yet be attempted. However, qualitative comparisons can be made with rather striking results.

Numerous road and dynamometer tests on four kinds of brake linings suffice to rank them in the following order:

1 Lining A—quiet under almost all conditions.
2 Lining B—normally quiet but occasionally noisy.
3 Lining C—similar to B, but definitely more prone to noise.
4 Lining D—too noisy under almost all conditions to be acceptable.

E_s and $(d\mu)/(dV)$ were measured for these four linings, as reported in the companion paper.[4] For reference, the results are repeated in Table 1.

According to the theory, lining A should be far less subject to vibration than the others, first, because the rate of build-up of vibration is less (low $d\mu/dV$), and also, because the low E_s corresponds to a natural frequency too low to induce resonance.

The other three show a rather closely spaced set of $(d\mu)/(dV)$ values (the ratios are 11.57:10:10.63), suggesting strongly that, for this group, the resonance term will be of decisive importance in determining the tendency to vibration; that is, the probability of noise will be ranked in the same order as E_s, with the highest E_s corresponding to the noisiest lining. Exactly this is observed.

The correctness of the theory obviously is not to be established on the basis of the present evidence alone. Time and additional data will be necessary. It is true, nevertheless, that the theory is not in disagreement with any known facts, and that it is capable of correct qualitative predictions. Whatever modifications may be necessary later, it should serve a useful purpose as a first attempt to introduce order into a very confused field, and as a guide for future experiments.

Bibliography

1 "Frictional Relaxation Oscillations," by B. R. Dudley and H. W. Swift, *Philosophical Magazine*, vol. 40, 1949, pp. 849–861.
2 "Frictional Vibrations," by D. Sinclair, *Journal of Applied Mechanics*, Trans. ASME, vol. 77, 1955, pp. 207–214.
3 "The Nature of the Static and Kinetic Coefficients of Friction," by E. Rabinowicz, *Journal of Applied Physics*, vol. 22, 1951, p. 1373.
4 "The Statistical Nature of Friction," by E. Rabinowicz, B. G. Rightmire, C. E. Tedholm, and R. E. Williams, Trans. ASME, vol. 77, 1955, pp. 981–984.
5 "Autocorrelation Analysis of the Sliding Process," by E. Rabinowicz, *Journal of Applied Physics*, vol. 27, 1956, pp. 131–135.

Self-Excited Vibrations of an Air-Lubricated Thrust Bearing

By L. LICHT,[1] D. D. FULLER,[2] and B. STERNLICHT[3]

Nomenclature

The following nomenclature is used in the paper:

A = area, in.^{2}
C = constant
H = annulus height, in.
$h = H - H_0$, small deviation from equilibrium in annulus height, in.
M = mass of air between bearing plates, $\frac{\text{lb-sec}^2}{\text{in.}}$
m = mass of upper plate (including load), $\frac{\text{lb-sec}^2}{\text{in.}}$
P = pressure, psia
$p = P_r - P_0$, small deviation from equilibrium of recess pressure, psi
$\mathbf{R}$ = gas constant, $\frac{\text{in.}^2}{\text{sec}^2\text{-deg R}}$
R = outer radius of bearing, in.
r = radius, in.
$q = \left(\frac{\partial M}{\partial P}\right)_0$ variation of air mass in bearing with respect to pressure at equilibrium, in-sec^2
$s = \left(\frac{\partial M}{\partial H}\right)_0$ variation of air mass in bearing with respect to annulus height at equilibrium, $\frac{\text{lb-sec}^2}{\text{in.}^2}$
T = temperature, deg R
t = time, sec
W = rate of mass flow, $\frac{\text{lb-sec}}{\text{in.}}$
$w = w_1 - w_2$ difference in mass flow rates, $\frac{\text{lb-sec}}{\text{in.}}$
$\alpha = \left(\frac{dW_1}{dP}\right)_0$ variation of mass-flow rate into the bearing with recess pressure at equilibrium, in-sec
$\beta = \left(\frac{\partial W_2}{\partial P}\right)_0$ variation of mass-flow rate out of the bearing with recess pressure at equilibrium, in-sec
$\theta = \left(\frac{\partial W_2}{\partial H}\right)_0$ variation of mass-flow rate out of the bearing with annulus height at equilibrium, $\frac{\text{lb-sec}}{\text{in.}^2}$
γ = ratio of specific heats C_p/C_v
Δ = depth of recess, in.
ρ = air density, $\frac{\text{lb-sec}^2}{\text{in.}^4}$

Subscripts

1 = into bearing
2 = out of bearing
a = annulus
at = atmosphere
c = critical
e = effective
0 = equilibrium condition
r = recess
s = supply
0, 1, 2 = in Equations [14] and [15] refer to constants
1, 2, e = in Fig. 4 refers to small diameter nozzle, large diameter nozzle, and capillary, respectively

Superscripts

. = first derivative
.. = second derivative
... = third derivative
} with respect to time

Introduction

The present trend toward high speeds, high temperatures, radioactive atmospheres, and low frictional requirements has renewed interest in gas-lubricated bearings. There are several very important factors to be considered before compressible fluids can be employed efficiently in bearing design. This paper does not concern itself with the compressor power consumption or bearing load-carrying capacity, but rather with the troublesome problem of instability. Several papers, references (1, 2, 3), make reference to this phenomenon without, however, attempting to analyze it. The authors in this paper present an approach to the investigation of the stability of an air-lubricated thrust bearing.

The stability analysis is based on a number of simplifying assumptions. It is not intended to provide definite design parameters, but rather to indicate the primary causes of the undesirable "air-hammer" phenomenon, to establish stability criteria, and to point out the parameters which influence it.

Bearing Configuration

The bearing consists of two circular plates, Figs. 1 and 2. The upper plate has a circular recess (recess diameter is $^1/_2$ of the bearing outside diameter), the depth of which can be varied at will. This plate supports the load, (mass m).

Air enters the recess from a constant-pressure reservoir via a bell-mouthed nozzle situated in the center of the lower plate. From there it flows through the narrow annulus (height H_0) radially outward to the atmosphere. The load which can be supported depends on the pressure maintained in the recess and the annulus.

Assumptions

At equilibrium the recess area is subjected to a uniform pressure P_0. The pressure drop from the edge of the recess to the bearing periphery is sensibly linear, references (4 and 5). It is assumed that for small deviations from the equilibrium point, this type of

[1] Columbia University, New York, N. Y. Assoc. Mem. ASME.
[2] Professor, Department of Mechanical Engineering, Columbia University, New York, N. Y. Mem. ASME.
[3] General Engineering Laboratory, General Electric Company, Schenectady, N. Y. Mem. ASME.
Contributed by the Lubrication Division and presented at the ASLE-ASME Lubrication Conference, Toronto, Ontario, Canada, October 7–9, 1957.
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 10, 1957. Paper No. 57—LUB-2.

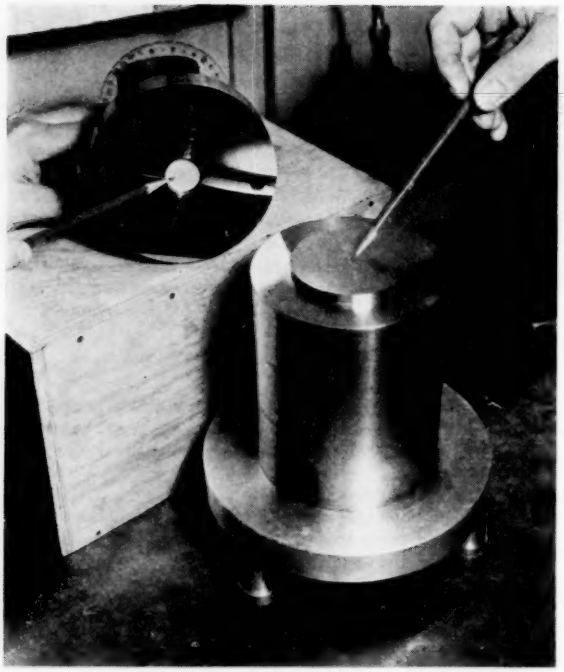


Fig. 1 Bearing configuration

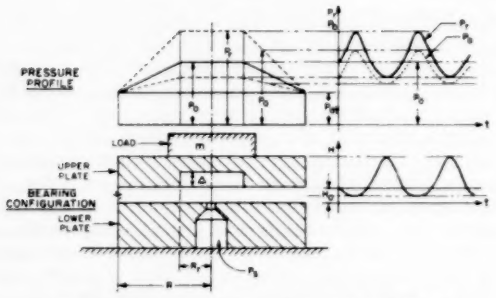


Fig. 2 Diagrammatic representation of oscillations

pressure distribution is preserved. Otherwise stated, the instantaneous pressure distribution can be represented by a frustum of height $(P_0 - P_{at}) + p$ when the corresponding annulus height is $H_0 + h$, p and h representing small deviations from the equilibrium values, Fig. 2.

Changes in air density are attributed mainly to the variation of pressure. For simplicity, the relationship $P/\rho = \mathbf{R}\ T_0$ will be considered to hold throughout, but a similar analysis can be performed on the basis of a polytropic relationship $p/\rho^n = \text{const}$.

Equations of Motion

Following the assumptions made, the pressure in the annulus is

$$P_a = P_r - (P_r - P_{at})\ \frac{r - R_r}{R - R_r} \quad \ldots\ldots\ldots\ldots [1]$$

Neglecting external damping, and considering that the upper plate is constrained to move in the vertical direction only, the equation of motion as shown in Appendix 1, is

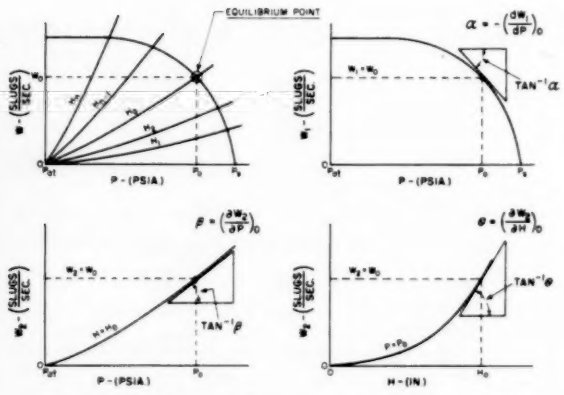


Fig. 3 Rates of change of mass flow about equilibrium point

$$m\ddot{h} = pA_e \quad \ldots\ldots\ldots\ldots\ldots\ldots [2]$$

where

$$\left.\begin{aligned} A_e &= \pi R_e^2 \\ R_e^2 &= R^2 - \frac{2R^3 + R_r^3 - 3R^2R_r}{3(R - R_r)} \end{aligned}\right\} \quad \ldots\ldots\ldots [3]$$

Referring to Fig. 3, it is noted that the mass flow into the bearing depends on the recess pressure only, whereas the outflow is a function of the recess pressure, as well as the annulus height.

To small deviations from the equilibrium point (p and h) there correspond variations in inflow and outflow which, to the first degree of approximation, can be written respectively as

$$\left.\begin{aligned} w_1 &= \left(\frac{dW_1}{dP}\right)_0 p = -\alpha p \\ w_2 &= \left(\frac{\partial W_2}{\partial P}\right)_0 p + \left(\frac{\partial W_2}{\partial H}\right)_0 h = \beta p + \theta h \end{aligned}\right\} \quad \ldots\ldots [4]$$

The time rate of change of the bearing air mass content then becomes

$$w = w_1 - w_2 = -(\alpha + \beta)\ p - \theta h \quad \ldots\ldots\ldots [5]$$

where α, β, θ, are all positive, Fig. 3.

The air mass contained between the bearing surfaces is

$$M = 2\pi \left\{\int_0^{R_r} (\Delta + H)\rho_r r dr + \int_{R_r}^{R} H\rho_a r dr\right\} \quad \ldots\ldots [6]$$

Using Equations [1] and [3] this expression reduces to (see Appendix 2)

$$M = \frac{1}{\mathbf{R}T_0}\ [HP_rA_e + \Delta P_r\pi R_r^2 + HP_{at}(\pi R^2 - A_e)] \ldots [7]$$

The time rate of change of the bearing air content $\dot{M}$ is evidently equal to the difference between inflow and outflow w and corresponds to the time rates of small deviations from the equilibrium point $\dot{p}$ and $\dot{h}$

$$w = \dot{M} = \left(\frac{\partial M}{\partial P}\right)_0 \dot{p} + \left(\frac{\partial M}{\partial H}\right)_0 \dot{h} = q\dot{p} + s\dot{h} \quad \ldots\ldots [8]$$

where, by differentiation of Equation [7]

$$q = \left(\frac{\partial M}{\partial P}\right)_0 = \frac{A_eH_0 + \Delta\pi R_r^2}{\mathbf{R}T_0}$$

$$s = \left(\frac{\partial M}{\partial H}\right)_0 = \frac{A_e(P_0 - P_{at}) + \pi R^2 P_{at}}{RT_0} \quad \ldots\ldots [9]$$

and

$$\frac{q}{s} = \frac{A_s H_0 + \Delta \pi R_r^2}{A_e(P_0 - P_{at}) + \pi R^2 P_{at}} \quad \ldots\ldots [10]$$

From Equations [5] and [8] we have

$$q\dot{p} + s\dot{h} + (\alpha + \beta)\, p + \theta h = 0 \quad \ldots\ldots [11]$$

and from Equation [2]

$$p = \frac{m}{A_e}\ddot{h}$$

$$\dot{p} = \frac{m}{A_e}\dddot{h} \quad \ldots\ldots [12]$$

Eliminating p and $\dot{p}$ between Equations [11] and [12], the following differential equation is obtained

$$\dddot{h} + \frac{\alpha + \beta}{q}\ddot{h} + \frac{sA_e}{mq}\dot{h} + \frac{\theta A_e}{mq}h = 0 \quad \ldots\ldots [13]$$

Stability Criteria

Equation [13] is of the form

$$\dddot{h} + C_2\ddot{h} + C_1\dot{h} + C_0 h = 0 \quad \ldots\ldots [14]$$

where all coefficients C are positive.

Applying Routh's stability criteria, reference (6), to Equation [14], the following inequality must be satisfied in order to achieve stability

$$C_1 C_2 > C_0 \quad \ldots\ldots [15]$$

where

$$C_2 = \frac{\alpha + \beta}{q}$$

$$C_1 = \frac{sA_e}{mq}$$

$$C_0 = \frac{\theta A_e}{mq}$$

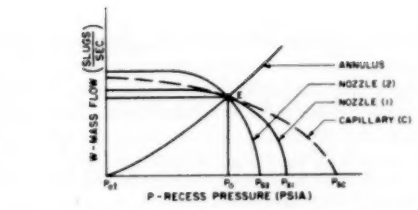
Therefore, stability criteria may be represented by the inequality

$$\frac{\alpha + \beta}{\theta} > \frac{q}{s} \quad \ldots\ldots [16]$$

It is of interest to discuss the effect of the various parameters on stability. A study of Fig. 3 indicates that large values of the ratio $\alpha + \beta/\theta$ correspond to large values of the recess pressure P_0 and small values of the annulus height H_0. For a given supply pressure P_s, a favorable condition results if the maximum possible load is being supported within the safety limits of a minimum annular height H_0. Under those conditions the ratio of α/θ has a large value, though, unavoidably, β is small. Equation [10] shows that the value of the q/s is proportional to the recess depth Δ, the annulus height H_0, and inversely to the recess pressure P_0. It can thus be noted that the values of P_0 and H_0 have an opposite effect on the magnitudes of the ratios forming the two sides of the Inequality [16].

It is also clear from Equation [10] that the recess, which represents the bulk of the air-storage capacity, should have a minimum depth Δ in order to achieve stability.

Referring to Fig. 4, it is apparent that the magnitude of α depends on the manner in which the air is supplied to the bearing. The three values of α correspond to conditions when the air is fed through a small nozzle α_1, a larger nozzle α_2, and a capillary α_c. In each case, the load is the same since the recess pressure P_0, but not the supply pressure P_s, and annular height H_0 remain unchanged. For a constant supply pressure the load determines uniquely the annulus height and recess pressure. Limiting values of the recess depth Δ_c were obtained for various loads for which a given pressure P_0 had to be maintained in the recess. Experimental results for a supply pressure of 73.5 psig are shown in Fig. 5 and indicate the effect of the recess depth, recess pressure, and nozzle size on stability. The influence of these parameters agrees with the trend predicted by the Inequality [16].



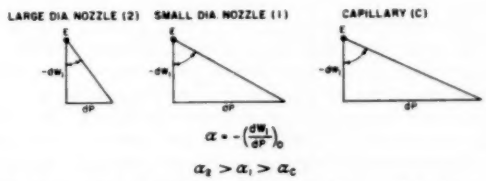


Fig. 4 Comparison of magnitudes of the coefficient α

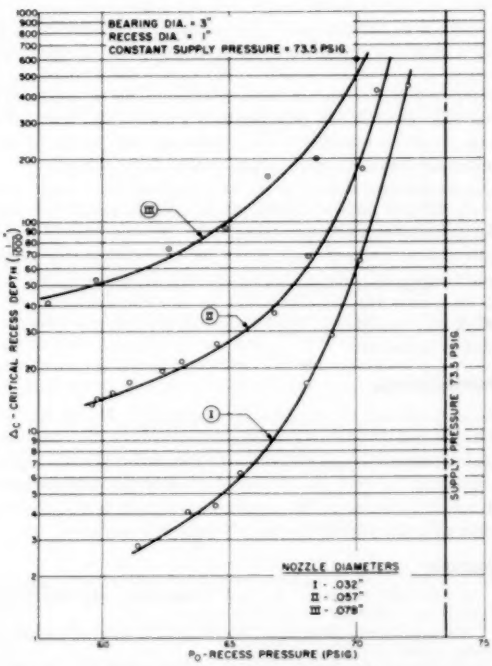


Fig. 5 Experimental results, critical recess depth versus recess pressure

Conclusions

The instability phenomenon is one of the primary considerations in the design of externally pressurized, gas-lubricated bearings. The proportions into which a bearing surface is subdivided, namely recess and annulus areas, is dictated by considerations of air consumption and load-carrying capacity. Stability, however, requires examination of the bearing geometry from the point of view of air-storage capacity. This should be held at a minimum. Consequently the recess depth Δ should be small in order to achieve stability. Generally, this depth should be comparable in magnitude to the annulus height. It is, by far, the most important parameter.

In all cases, it is desirable that the difference between the supply pressure P_s and recess pressure P_r be small. A condition wherein the nozzle is choked will cause the bearing to be unstable, unless the recess depth is made very small. The nozzle size is limited since it functions as a flow restrictor, but within these limitations the largest possible nozzle diameter should be used. The validity of the above considerations is substantiated by the experimental results shown in Fig. 5.

The inferiority of capillaries as restrictors can best be illustrated by an experimental result. When a 220-in-long, 0.032-in-diam tube was substituted for a nozzle of the same diameter, stability could only be achieved at the expense of a considerable reduction in the recess depth. Moreover, the volume of the capillary represents an air storage capacitance in addition to that of the recess which has already been shown to have a very adverse effect on stability. The observed frequencies of these self-excited vibrations were of the order of 25 to 30 cps.

Adverse combinations of low recess pressures and large recess depths produced double amplitudes of oscillations well in excess of the annular equilibrium height H_0. At the same time, no metal-to-metal contact took place on the downstroke despite the narrowness of the annulus separating the plates 0.002 in. to 0.003 in. This can be attributed to the "squeeze action" of the fluid film resulting in the increase of the reactive force when the width decreases.

Acknowledgments

The authors wish to express their gratitude to Mr. H. Apkarian of the General Electric Company, Schenectady, for his assistance and numerous suggestions in this study. The authors are also indebted to Prof. J. P. DenHartog of the Massachusetts Institute of Technology for suggesting a simple approach to the problem and the interest he has shown in this work.

Bibliography

1 "Air Driven Spinners," by L. E. Wightman, *Machine Design*, May, 1948, pp. 121–125.

2 "Air Lubricated Bearings," by P. M. Mueller, *Product Engineering*, August, 1951, pp. 112–115.

3 "Air Bearing Studies at Normal and Elevated Temperatures," by J. D. Pigott and E. F. Macks, *Lubrication Engineering*, February, 1954, pp. 29–33.

4 "Preliminary Investigation of an Air Lubricated Hydrostatic Thrust Bearing," by L. Licht and D. D. Fuller, ASME Paper No. 54—Lub-18.

5 "Temperature Effects in Hydrostatic Thrust Bearing Lubrication," by W. F. Hughes and J. F. Osterle, ASME Paper No. 55—Lub-11.

6 "Mechanical Vibrations," by J. P. DenHartog, McGraw-Hill Publishing Company, Inc., New York, N. Y., third edition, 1947.

APPENDIX 1

Assuming a linear pressure gradient in the annulus Fig. 2, Equation [2] can be written as

$$\begin{aligned} m\ddot{k} &= 2\pi\left[\int_0^R prdr - \int_{R_r}^R p\,\frac{r-R_r}{R-R_r}\,rdr\right] \\ &= p\pi\left[R^2 - \frac{2R^3 + R_r^3 - 3R^2R_r}{3(R-R_r)}\right] \\ &= p\pi R_e^2 \\ &= pA_e \end{aligned}$$

where

$$R_e^2 = R^2 - \frac{2R^3 + R_r^3 - 3R^2R_r}{3(R-R_r)}$$

$$A_e = \pi R_e^2$$

APPENDIX 2

The air-mass content of the bearing is:

$$M = 2\pi\left[\int_0^{R_r}(\Delta + H)\rho_r r dr + \int_{R_r}^R H\,\rho_a\, r\,dr\right]$$

Assuming $\rho_r = \dfrac{P_r}{\mathbf{R}T_0}$ and making use of Equation [1]

$$\rho_a = \frac{P_a}{\mathbf{R}T_0} = \frac{1}{\mathbf{R}T_0}\left[P_r - (P_r - P_{at})\,\frac{r-R_r}{R-R_r}\right]$$

substitution into the integrants gives

$$\begin{aligned} M &= \frac{2\pi}{\mathbf{R}T_0}\left[HP_r\int_0^R rdr + \Delta\cdot P_r\int_0^{R_r} rdr \right. \\ &\qquad \left. - H\,\frac{P_r - P_a}{R-R_r}\int_{R_r}^R (r-R_r)\,rdr\right] \\ &= \frac{\pi}{\mathbf{R}T_0}\left\{HP_r\left[R^2 - \frac{2R^3 + R_r^3 - 3R^2R_r}{3(R-R_r)}\right]\right. \\ &\qquad \left. + \Delta\cdot P_r R^2 + HP_{at}\left[\frac{2R^3 + R_r^3 - 3R^2R_r}{3\,(R-R_r)}\right]\right\} \end{aligned}$$

$$M = \frac{1}{\mathbf{R}T_0}\left[HP_rA_e + \Delta\cdot P_r\pi R_r^2 + HP_{at}\,(\pi R^2 - A_e)\right]$$

where R_e and A_e are as defined in Appendix 1.

A Simple Formula for Determining the Position of Maximum Slider Velocity in a Slider-Crank Mechanism

By CHING-U IP[1] and L. C. PRICE,[2] EAST LANSING, MICH.

A cubic equation which gives the position of maximum slider velocity is derived. The equation lends itself readily to be solved to any desired degree of accuracy by Lin's method. A simple formula is found to furnish a closed-form answer which is accurate within 4 min of a degree for *l/r* ratio of 1.5, and has practically no error for *l/r* greater than 5. The results are compared with those obtained from the familiar approximate slider-velocity formula having a second harmonic.

The Approximate Solution

THE familiar approximate formulas for the velocity and the acceleration of the slider in a slider-crank mechanism, Fig. 1, are respectively

$$V = \omega r \left(\sin\theta + \frac{1}{2n} \sin 2\theta \right) \quad \ldots\ldots [1]$$

$$A = \omega^2 r \left(\cos\theta + \frac{1}{n} \cos 2\theta \right) \quad \ldots\ldots [2]$$

where $\omega = d\theta/dt$ = angular velocity of crank
$n = l/r$ = connecting-rod-to-crank ratio
θ = crank angle measured from top dead center

For maximum velocity, the acceleration will be zero, or

$$\cos\theta + \frac{1}{n}(2\cos^2\theta - 1) = 0$$

Substituting x for $\cos\theta$ gives the quadratic equation

$$2x^2 + nx - 1 = 0 \quad \ldots\ldots [3]$$

the solution of which is

$$\cos\theta = x = \frac{-n}{4} + \left(\frac{n^2}{16} + \frac{1}{2} \right)^{1/2} \quad \ldots\ldots [4]$$

The + sign before the radical is chosen because the absolute value of $\cos\theta$ cannot be greater than unity.

The Characteristic Cubic Equation

The exact formulas[3] for the velocity and the acceleration of the slider are, respectively

$$V = \omega r \left[\sin\theta + \frac{\sin 2\theta}{2(n^2 - \sin^2\theta)^{1/2}} \right] \quad \ldots\ldots [5]$$

$$A = \omega^2 r \left[\cos\theta + \frac{n^2 \cos 2\theta + \sin^4\theta}{(n^2 - \sin^2\theta)^{3/2}} \right] \quad \ldots\ldots [6]$$

For maximum velocity, the acceleration will be zero, or

$$\cos\theta[n^2 - (1 - \cos^2\theta)]^{3/2} + n^2(2\cos^2\theta - 1) + (1 - \cos^2\theta)^2 = 0$$

Substituting x for $\cos\theta$ gives

$$x(n^2 - 1 + x^2)^{3/2} = (n^2 - 1) - 2(n^2 - 1)x^2 - x^4$$

Substituting $m = n^2 - 1$, $y = x^2 = \cos^2\theta$, and squaring both sides of the equation and simplifying, give the characteristic cubic equation

$$y^3 + (m - 2)y^2 - (m^2 + 4m)y + m = 0 \quad \ldots\ldots [7]$$

Descarte's rule of signs indicates that this equation has two positive real roots. However, we are only interested in the positive root which is less than or equal to unity.

Lin's Method[4] of Solving the Equation

Lin's method is a successive synthetic division method and is so simple that it can be memorized. Furthermore, it can be made as accurate as occasion demands. Consider the case of $n = l/r = 3$, or, $m = n^2 - 1 = 8$; the characteristic equation will be

$$y^3 + 6y^2 - 96y + 8 = 0 \quad \ldots\ldots [8]$$

Lin's method forms the first trial divisor from the last two terms of the polynomial; thus

$$y - \frac{8}{96} \quad \text{or} \quad y - 0.0834$$

The long-division process can be conducted in synthetic form

$$\begin{array}{r|l} & y^2 + 6.0834y - 95.493 \\ \hline y - 0.0834 & y^3 + 6y^2 - 96y + 8 \\ & y^3 - 0.834y^2 \\ \hline & 6.0834y^2 - 96y \\ & 6.0834y^2 - 0.507y \\ \hline & -95.493y + 8 \\ & -95.493y + 7.98 \\ \hline \end{array}$$

Remainder not quite zero

The second trial divisor is formed from

$$y - \frac{8}{95.493}, \quad \text{or} \quad y - 0.0838$$

---

[1] Assistant Professor in Mechanical Engineering, Michigan State University. Assoc. Mem. ASME.

[2] Professor and Head of Mechanical Engineering Department, Michigan State University. Mem. ASME.

[3] See "Theory of Machines," by T. Bevan, Longmans, Green & Company, London, England, 1939, p. 84.

Contributed by the Machine Design Division and presented at the Spring Meeting, Birmingham, Ala., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 8, 1956. Paper No. 57—S-8.

[4] "Method of Successive Approximations of Evaluating the Real and Complex Roots of Cubic and Higher Order Equations," by S. N. Lin, *Journal of Mathematics and Physics*, vol. 20, August, 1941, pp. 231–242.

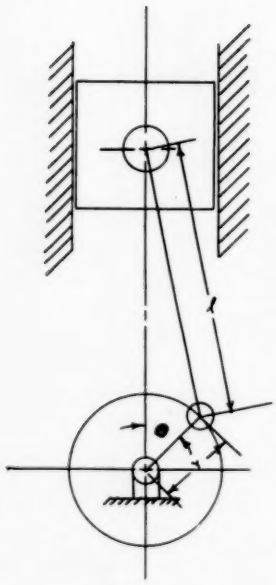


FIG. 1

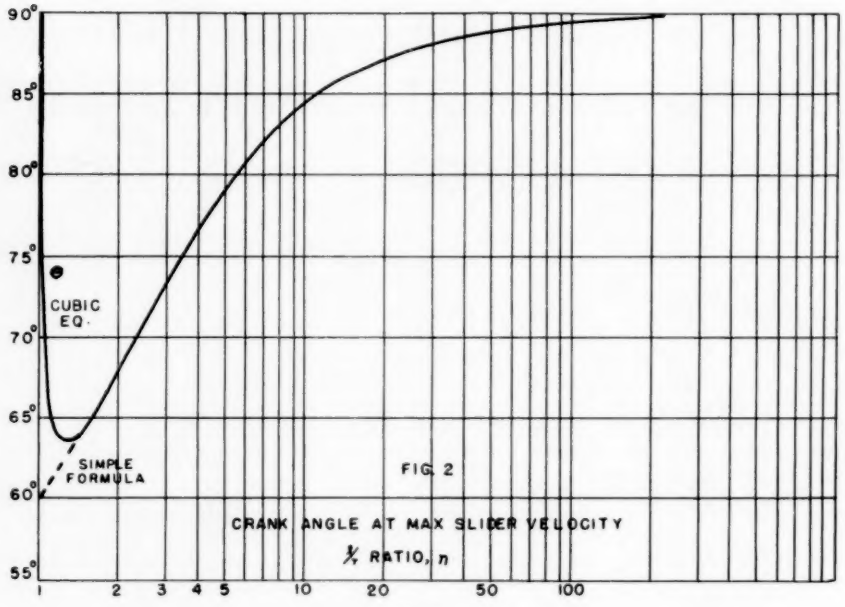


FIG. 2

TABLE 1 COMPARISON OF RESULTS OBTAINED FROM SIMPLE FORMULA WITH SOLUTION BY CUBIC EQUATION

	Solution of cubic equation		Simple formula $\cos\theta = (n^2+3)^{-1/2}$		
$n = l/r$	$\cos\theta$	θ	$\cos\theta$	θ	Difference
1	0	90°	0.5000	60°	+30°
1.01	0.2898	73° 9′	0.4988	60° 5′	+13° 4′
1.1	0.4231	64° 58′	0.4874	60° 50′	+ 4° 8′
1.3	0.4467	63° 28′	0.4546	62° 58′	+ 30′
1.5	0.4359	64° 10′	0.4364	64° 7′	+ 3′
2	0.3789	67° 45′	0.3779	67° 48′	− 3′
3	0.2895	73° 10′	0.2887	73° 13′	− 3′
4	0.2300	76° 42′	0.2294	76° 44′	− 2′
5	0.1892	79° 6′	0.1900	79° 7′	− 1′
6	0.1600	80° 47′	0.1601	80° 47′	0
7	0.1387	82° 2′	0.1387	82° 2′	0
8	0.1222	82° 59′	0.1222	82° 59′	0
9	0.1091	83° 44′	0.1091	83° 44′	0
10	0.0985	84° 21′	0.0985	84° 21′	0
15	0.0662	86° 12′	0.0662	86° 12′	0
20	0.0499	87° 8′	0.0499	87° 8′	0
50	0.0200	88° 51′	0.0200	88° 51′	0
100	0.0100	89° 26′	0.0100	89° 26′	0
∞	0	90°	0	90°	0

which gives substantially no remainder term after the long-division process. Otherwise, the process is continued until the remainder is zero. Therefore, the factors for Equation [8] are

$$(y - 0.0838)(y^2 + 6.0838y - 95.49) = 0 \quad \ldots\ldots [9]$$

The quadratic term gives roots greater than unity; therefore

$$\cos^2\theta = y = 0.0838$$

or

$$\cos\theta = 0.28948$$

$$\theta = 73° 10' \quad \text{or} \quad 286° 50' \quad \ldots\ldots [10]$$

THE SIMPLE CLOSED-FORM SOLUTION

The fact that the first trial divisor gives practically no remainder leads one to believe a very close approximation of Equation [7] to be

TABLE 2 RESULTS OF SOLUTION OBTAINED FROM APPROXIMATE SLIDER-VELOCITY FORMULA

	Solution of cubic equation		$\cos\theta = -\frac{n}{4} + \left(\frac{n^2}{16} + \frac{1}{2}\right)^{1/2}$		
$n = l/r$	$\cos\theta$	θ	$\cos\theta$	θ	Difference
1	0	90°	0.5000	60°	+30°
1.01	0.2898	73° 9′	0.4983	60° 7′	+13° 2′
1.1	0.4231	64° 58′	0.4837	61° 3′	+ 3° 55′
1.3	0.4467	63° 28′	0.4532	63° 3′	+ 25′
1.5	0.4359	64° 10′	0.4254	64° 49′	− 39′
2	0.3789	67° 45′	0.3660	68° 32′	− 47′
3	0.2895	73° 10′	0.2808	73° 42′	− 31′
4	0.2300	76° 42′	0.2247	77° 1′	− 19′
5	0.1892	79° 6′	0.1861	79° 17′	− 11′
6	0.1600	80° 47′	0.1583	80° 54′	− 7′
7	0.1387	82° 2′	0.1374	82° 6′	− 4′
8	0.1222	82° 59′	0.1213	83° 2′	− 3′
9	0.1091	83° 44′	0.1084	83° 46′	− 2′
10	0.0985	84° 21′	0.0981	84° 22′	− 1′
15	0.0662	86° 12′	0.0660	86° 13′	− 1′
20	0.0499	87° 8′	0.0498	87° 9′	− 1′
50	0.0200	88° 51′	0.0200	88° 51′	0
100	0.0100	89° 26′	0.0100	89° 26′	0
∞	0	90°	0	90°	0

$$(m^2 + 4m)y - m = 0 \quad \ldots\ldots [11]$$

the solution of which is

$$y = \frac{1}{m+4} = \frac{1}{n^2+3}$$

Hence

$$\cos\theta = y^{1/2} = (n^2+3)^{-1/2} \quad \ldots\ldots [12]$$

It is a well-known fact that if n is infinite then θ will be 90 deg for the slider velocity to be a maximum. Both the characteristic cubic equation, Equation [7], and the simple formula Equation [12] give this result for the limiting case.

Accuracy of the Simple Formula

The accuracy of the results obtained from the simple formula is investigated by comparing with those obtained through the solution of the cubic equation. The comparison is shown in Table 1, where cos θ is carried to the fourth place of decimals, and θ is recorded to the nearest minute. The two solutions (in the form of θ versus n curves) are plotted in semi-logarithmic co-ordinates in Fig. 2.

Accuracy of the "Approximate Solution"

The accuracy of the solution, obtained from the approximate slider-velocity formula having a second harmonic, is similarly investigated. The results are entered in Table 2.

Conclusions

(*a*) In the tabulated results, it is understood that if θ is a solution, then 360 deg $- \theta$ is also a solution.

(*b*) In the region of n greater than 1.5, the "simple formula" is more accurate than the "approximate solution," and is simpler to use.

(*c*) In the region of $1.0 \leq n < 1.5$, only the solution of the cubic equation gives the accurate result.

Discussion

A. E. Richard de Jonge.[5] The subject matter discussed by the authors has been dealt with exhaustively as far back as 1896 and 1898, that is, about 60 years ago, in the English technical literature. At that time, the cubic equation was derived and tables of values for different ratios $n = 1/r$ were given by several authors.

As proof, a number of references are cited[6,7,8] which the authors could have found easily had they done but a little searching in the literature before rushing into print.

All these give the cubic equation, but instead of in the cosine form, in the sine form which provides simpler coefficients. Burls gives an extensive table of even closer values than those of the authors, and if the values of the angle θ for the various n-values would not have been correct, they would have been challenged probably at the time they were published. It appears, therefore, to be necessary to check the authors' values as they differ appreciably from the values as given, for example, by Burls especially in the range from $n = 7$ to $n = 10$. Consequently, it looks as if the authors' approximate formula gives very accurate values which apparently it does not.

The cubic equation by Hill-Unwin-Burls also has been given in a number of English textbooks together with tables of θ for various values of n as for example by Low[9] and McKay.[10]

In addition, there is an even earlier German reference by Schadwill, which was cited in the early editions of "Die Hütte, des Ingenieurs Taschenbuch."[11] This presents a somewhat different cubic equation for the solution of the crank angle for maximum piston velocity.

Relatively recently, the cubic equation as given by Unwin was published anew by Freudenstein (his Equation [45]).[12] He, too, probably was not aware of its previous existence.

Thus, it appears, that the only thing new produced by the authors is the approximate equation obtained from Lin's method of solving the cubic equation. However, this suffers from the same fault as do earlier approximate equations as published by Unwin and others, in that the values obtained from them give close approximation for a very limited range of n-values only, but differ widely from the true values in other ranges. Inasmuch as the accurate angles for maximum crosshead velocity and n-values from $n = 1$ to $n = 10$, and $n = \infty$ have been calculated and are known from tables published, as for instance by Burls, there seems to be no valid reason for trying to derive further approximate equations which also give near accurate values over a very small range of n-values only.

In addition, even the curves of the diagram given by the authors are similar to those published by Burls.

Ferdinand Freudenstein.[13] In connection with the authors' elegant short formula, mention may be made of corroborative results of another investigation,[12] in which Equation [45] (page 785) and Fig. 10 (page 784) correspond to the authors' Equation [7] and Fig. 2, respectively.

W. F. Vogel.[14] The exact solution of the problem of maximum slider velocity in a slider-crank mechanism has been attempted in several publications. The "characteristic cubic equation" of the authors has been known for many decades.

This newest attempt, like most of its predecessors, failed to come up with the perfect answer, because the authors did not believe in the possibility of a closed-term solution of the cubic equation. Therefore, they resorted to an approximate formula, derived from trial-and-error solutions of the cubic equation.

Actually, a closed-term solution of the exact cubic equation is not only possible, but its algebraic expressions are surprisingly simple. This solution has been given in the third installment of a series of articles by the writer.[15]

The analytical solution is so complete that even a graphical construction for the maximum velocity of the slider and for the corresponding position of all members of the linkage could be presented in the same publication. Also included are sufficiently precise nomographs, from which numerical values can be read covering all of these details for any connecting-rod ratio.

The success in finding the aforementioned solution can be attributed to the discovery by the writer of a "special fixed point" of the mechanism, which enabled him to write the *general equations of its motion in expressions of unprecedented simplicity*. This discovery could not have been made without the treasure of knowledge found in other publications, which are quoted in the paper.[15]

The position of maximum slider velocity coincides with that of zero slider acceleration. This fact is of major importance,

[5] Mechanical Engineer, The Reeves Instrument Corporation, New York, N. Y., and Adjunct Professor, Polytechnic Institute of Brooklyn, Brooklyn, N. Y. Mem. ASME.

[6] "The Problem of the Connecting Rod," by M. J. M. Hill, Proceedings of the Institution of Civil Engineers, vol. 124, 1896, pp. 390–401.

[7] "Determination of Crank Angle for Greatest Piston Velocity," by W. C. Unwin, Proceedings of the Institution of Civil Engineers, vol. 125, 1896, pp. 363–366.

[8] "Note on Maximum Crosshead Velocity," by G. A. Burls, Proceedings of the Institution of Civil Engineers, vol. 131, 1898, pp. 338–346.

[9] "Applied Mechanics," by D. A. Low, Longmans, Green & Company, London, England, first edition, 1909, p. 304; second edition, 1913, p. 304.

[10] "The Theory of Machines," by R. F. McKay, Edward Arnold, London, England, 1915, p. 145.

[11] "Die Hütte, des Ingenieurs Taschenbuch," 19th edition, vol. 1, 1905, reference to Schadwill, p. 719.

[12] "On the Maximum and Minimum Velocities and the Accelerations in Four-Link Mechanisms," by F. Freudenstein, Trans. ASME, vol. 78, 1956, pp. 779–787.

[13] Associate Professor, Department of Mechanical Engineering, Columbia University, New York, N. Y. Assoc. Mem. ASME.

[14] Professor of Engineering Mechanics, Wayne State University, Detroit, Mich.

[15] "Crank Mechanism Motions—New Methods for Their Exact Determination," by W. F. Vogel, *Product Engineering*, vol. 12, 1941, pp. 423–428.

because it influences decisively the shape of the slider's acceleration-displacement curve, for which further details and graphical constructions were revealed in the publication cited.[15]

The new approximate formula of the authors for position of maximum slider velocity is very simple and superior in accuracy to any pertaining approximation the writer has seen published (including those of his own). Its errors are negligible in the range of the most frequent applications of the mechanism; i.e., in combustion and steam engines.

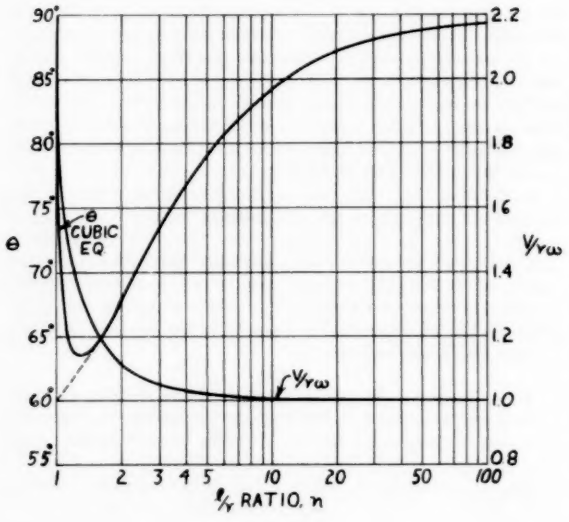


FIG. 3

W. C. TRIFTSHOUSER[16] AND A. S. HALL.[17] To the writers, the authors' simple and accurate expression for position of maximum slider velocity is a very interesting and surprising result. It encourages us to believe that simple forms may be found for solutions to other apparently complicated problems in kinematics.

As a matter of interest, the writers have superimposed on the authors' Fig. 2 a curve showing the ratio, $V/r\omega$, of maximum slider velocity to crankpin velocity. The result is as shown in Fig. 3 of this discussion.

AUTHORS' CLOSURE

The authors realize the topic is a classical problem that has been dealt with previously. Schadwill in a thesis presented in 1876 (not 1905), and entitled "Das Gliedervierseit," proved that the configuration of the slider-crank giving maximum velocity of the slider existed when the line of instantaneous centers is perpendicular to the connecting rod's direction. However, given a $1/r$ ratio this position cannot be graphically constructed. Klein[18] derived from Schadwill's proposition the cubic equation which was the same as that presented by Dr. Freudenstein and similar to that of the authors.

The authors did not, however, believe that a simple closed-form solution to the cubic equation was possible and were quite surprised to find that their simple approximate solution works for a very wide range of n. Dr. Vogel's monumental work certainly covered the subject of slider-crank motion thoroughly. His is the only exact solution the authors know, and whose existence the authors did not know previously.

The $V/r\omega$ curve of Mr. Triftshouser and Dr. Hall extends the usefulness of the result of the paper.

---

[16] Graduate Student, School of Mechanical Engineering, Purdue University, Lafayette, Ind.

[17] Professor of Mechanical Engineering, Purdue University, Lafayette, Ind. Mem. ASME.

[18] "High Speed Engine," by J. F. Klein, Van Nostrand and Company, New York, N. Y., 1911, Appendix A.

Some Methods for the Structural Design of Wings for Application Either at Ambient or Elevated Temperatures

By J. W. SEMONIAN[1] and R. F. CRAWFORD,[2] SANTA BARBARA, CALIF.

Some new methods are established for the design of aircraft-wing structures. The methods apply either at ambient or moderately elevated temperatures. Designs of a variety of structural configurations based on these methods are compared and discussed.

Nomenclature

The following nomenclature is used in the paper:

b_S = plate width between web supports
C = constant
D_x = flexural stiffness of compression cover in spanwise direction
d = over-all depth of beam
E = Young's modulus of elasticity
F = Farrar's efficiency factor
h = depth of internal structure
L = column length between rib supports
M_i = bending moment per unit width of beam chord
P_i = compressive loading per unit width of beam chord
t = thickness of compression plate
Σ = solidity of box beam
Σ_w = that fraction of box-beam volume occupied by web material
σ = stress
σ_{max} = maximum compressive stress
σ_{cy} = compressive yield stress
ρ = material density
η = plasticity reduction factor

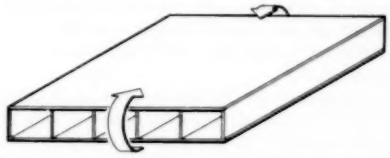
Introduction

Thermal flight and the associated increase in structural weight of the aircraft have been studied by many current investigators. Consequently, the lack of information on the efficient design of built-up structures has become increasingly apparent. Methods for the efficient design of box beams are therefore submitted as a contribution toward eliminating this deficiency.

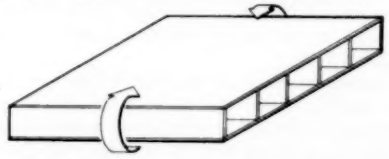
Box beams represent the primary structure of wings and other lift surfaces. Since the bending load on these surfaces is usually the major factor influencing their design, effort is concentrated on that type of loading.

Three basic types of box beams are studied which are categorized according to the nature of the internal structure as shown in Fig. 1:

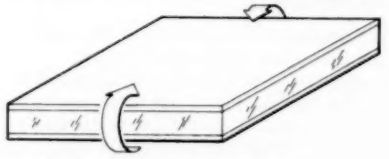
1 The various multiweb designs characterized by two covers connected by spanwise internal members at discrete spaces.



Group I – Spanwise Internal Members – Webs



Group II – Chordwise Internal Members – Ribs



Group III – Continuous Internal Medium

Fig. 1

2 Wide-column rib-supported configurations characterized by two covers connected by chordwise internal members at discrete spaces.

3 Two covers connected by a continuous internal medium.

These three basic groups include the more common configurations.

The most important effect of aerodynamic heating, the deterioration of mechanical properties of materials, is considered. However, in this study, thermal stresses resulting from differential expansion of the box-beam components are not emphasized because of a fortunate coincidence which became apparent in the investigation; namely, that design can proceed so that there is no conflict between designing for the aforementioned thermal stresses and for minimum weight.

With these restrictions, it was found possible to proceed with the development of criteria and methods for the efficient design of box beams.

In performing these studies, several original strength and efficiency analyses were made as necessary.

Analysis

Optimum design is defined here as that design which will accomplish a given function, usually stated in terms of strength, rigidity, temperature, time, and dimensions, for a minimum weight. In order to reduce the parameters and eliminate scale considerations, the loading-index concept, introduced by Za-

[1] Zahorski Engineering, Inc.
[2] Aerophysics Development Corporation. A Subsidiary of Curtiss-Wright Corporation.

Contributed by the Aviation Committee and presented at the Semi-Annual Meeting, San Francisco, California, June 9–13, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, June 11, 1957. This paper was not preprinted.

horski (1),[3] is used. The loading index used here for pure bending of a beam is in the form of the quotient of the moment per inch of chord and the square of the over-all depth of the beam. The loading index consolidates the specified load and dimensions in a parameter which has the units of stress. Gerard's solidity concept (2) is used as a means of expressing the efficiency of the built-up structure. Solidity is used here to represent the fraction of the cross-sectional area of a box beam occupied by structural material. The product of material density and solidity, which may be termed structural density, is used to compare beams of different materials on the same graph. The results of this investigation are presented as graphs with structural density as the ordinate and loading index as the abscissa. On this type of efficiency graph, high efficiency is associated with low structural density.

Aside from failures of the tension cover and joints and disregarding interaction, there exist three classes of failure due to bending which are applicable to all box-beam configurations considered here:

1 Compression failure of the cover panel without depthwise displacement of the internal structure (referred to here as the local mode of failure).

2 Compression failure of the internal structure due to flexure induced crushing forces.

3 General instability mode of failure characterized by depthwise displacement of the internal structure with chordwise nodes in the deflection pattern of the covers. This mode can occur with integral construction.

Internal structure here refers only to full-depth members connecting the tension and compression covers.

With recognition of the modes of failure which may occur and the assumption that there is no interaction among them, approximately optimum proportions of the structure may be determined. Use of methods similar to those of the efficiency analyses of Zahorski (1), Gerard (2), Shanley (3), and others is suggested; that is, for a given bending load the cover is designed to be critical in its local mode of failure; the internal structure will be of approximately optimum proportions if either:

1 Its strength is critical in the crushing mode of failure.

2 Or its stiffness is critical in the general instability mode of failure.

The internal structure must be designed both for strength and stiffness, the heavier design governing.

This design criterion does not mean that the structure should be critical in all three modes. That this would lead to nonoptimum proportions will be proved later.

The flat-web multiweb beam shown in Fig. 2 (one of the Group 1 configurations) has received much attention in the literature. This design was initially analyzed by Schuette and McCulloch (4) and has since been investigated by Gerard (2), Conway (5), Rosen (6), and others, using various failure criteria. Some of these investigators have used in part the foregoing approach; however, they differ somewhat on their choices of failure criteria. Some use buckling while others use various maximum-strength formulas. Efficiency curves have been computed here using the method of Rosen (6) and von Karman's equation (7) for the potential strength of the cover panels

$$\sigma = C(E\sigma_{cy})^{1/2}\left(\frac{t}{b_g}\right) \quad \ldots\ldots\ldots\ldots\ldots [1]$$

To demonstrate the methods, an example thermal-exposure condition of 400 F for 300 hr has been chosen and is used for all results presented here.

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

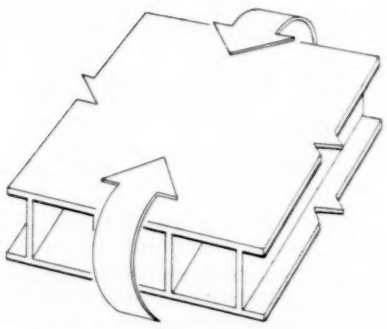


Fig. 2 Flat-plate multiweb beam

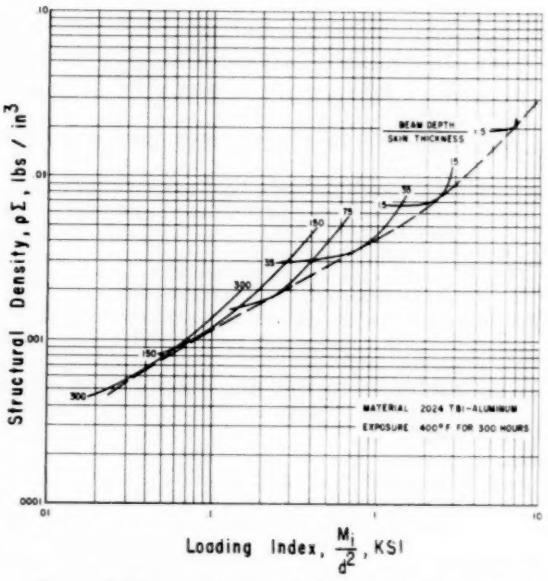


Fig. 3 Efficiency chart for flat-web multiweb beam

In Fig. 3 is shown a set of efficiency curves for multiweb beams of 2024-T81 aluminum alloy. Optimum efficiency is indicated by the lower envelope of the parametric curves. The parametric curves, of constant beam depth-to-skin thickness ratios may be useful when torsional-rigidity specifications require a certain skin thickness. These curves may be replotted with the ratio of beam depth-to-web spacing as a parameter to cover those cases for which the web spacing is specified. In Fig. 4 is shown a comparison of optimum efficiency curves for multiweb beams of 2024-T81 aluminum alloy, RC 130 A titanium alloy, and stainless W steel, each exposed to 400 F for 300 hr. It is seen, from this comparison, that, for the flat-web multiweb beam, the choice of materials for optimum design depends upon the loading index. Thus the choice of material ranges from aluminum to titanium to steel as loading index is increased.

The flat plate, in general, is not an efficient compression member. Therefore it cannot, in general, be expected that the most efficient box beam will be produced by the use of flat plates as webs and cover plates.

Investigation of suitable web and rib configurations led to the

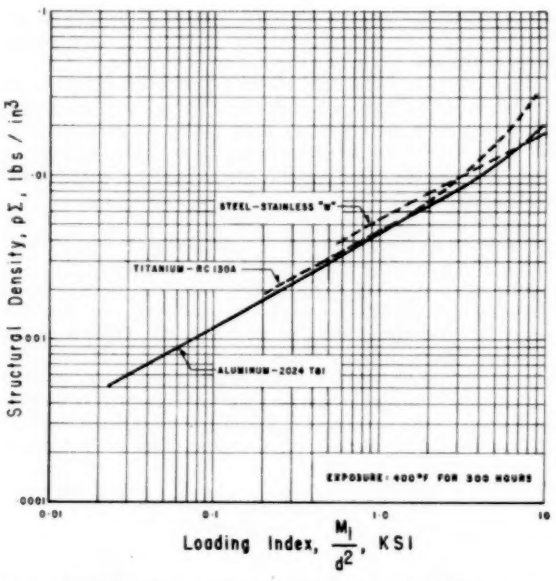


Fig. 4 Comparison of optimum designs for flat-web multiweb beams of three different materials

selection of corrugated sheet for application here. These configurations when optimized to carry the crushing loads were found to be more efficient than the optimized Z-stiffened panel if it were used as an internal member. Expressions for optimum design of corrugated compression members are obtained by applying the methods of reference (3).

In addition to being an efficient compression member, the corrugated web also carries shear loads efficiently, provides good fixity for the covers, and has negligible span-wise stiffness. This last property of the corrugated web is of particular importance in that it provides negligible restraint to thermal expansion of the covers, thereby virtually eliminating thermal stresses which arise through differential expansion of web or rib and cover.

In Fig. 5 is shown the mode of general instability for a box beam having corrugated-web internal structure. It is characterized by troughs and crests that extend across the entire width of the compression cover. The tension cover assumes a similar shape with lesser amplitude. This condition requires that the web be deformed compatibly.

Because no method appeared to be available to predict failure for the ideal and desirable case in which the webs are corrugated and integrally attached to the cover plates, an analysis was made. From an approximation of those results, the following simple equation is derived for web solidity required to prevent premature general instability

$$\Sigma_w = \left(\frac{P_i}{1.46\pi}\right)^2 \frac{h^2}{dD_x E} \quad \ldots\ldots\ldots\ldots\ldots\ldots [2]$$

For a given cover panel the design of integrally attached corrugated webs is executed by designing for flexure-induced crushing as explained previously and applying the foregoing formula for general instability, in the manner outlined. The corrugated-web multiweb beam, illustrated in Fig. 6, was investigated using von Karman's maximum panel-strength formula to predict failure of the compression cover.

In Fig. 7 is shown a set of efficiency curves for that design. As in the case of the flat-web multiweb beam, the lower envelope of the curves indicates optimum design. It is obvious that the designs for which the structure is critical in all three modes, indicated by the dotted line in Fig. 7, are not optimum.

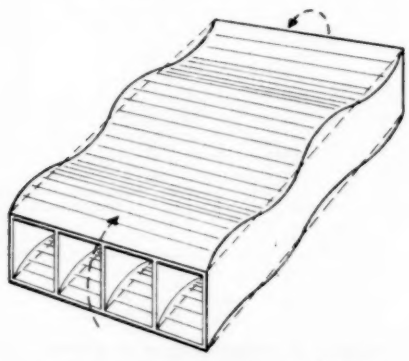


Fig. 5 Wrinkling or general instability

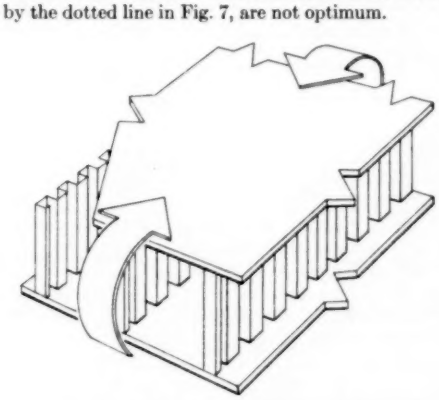


Fig. 6 Corrugated-web multiweb beam

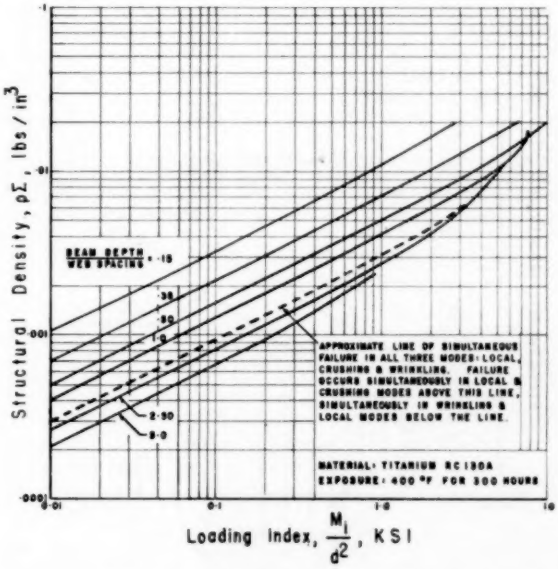


Fig. 7 Efficiency chart for multiweb beams with corrugated webs integral with covers; d/bs retained as design parameter

In Fig. 8 is shown a comparison of optimum designs for corrugated-web multiweb beams of 2024-T81 aluminum alloy, RC 130A titanium, and stainless W steel, each of which is exposed to 400 F for 300 hr. Again the trend shown in Fig. 4 is displayed; that is, the less dense material is more efficient in lower ranges of loading, while the higher density materials are more efficient in higher ranges of loading.

Fig. 9 shows a comparison of optimum designs for flat and corrugated-web multiweb beams. The optimum materials according to loading range are used in these envelopes; hence material varies along the envelope. Over the range of loading shown, the corrugated-web designs are significantly more efficient than the flat-web designs.

In addition to being very efficient in resisting bending, the corrugated-web design, as explained previously, does not permit significant thermal stresses to develop. This is not the case with flat-plate webs where, depending upon the rate of heating, serious problems can arise which would require addition of material and lead to reduced efficiency.

As discussed previously, the flat plate is in general an inefficient compression member; therefore the stiffened plate as illustrated in Fig. 10 was considered as a cover for the multiweb beam. The results of an efficiency study of this panel and the previously discussed methods for corrugated-web design were used to obtain the optimum efficiency chart shown in Fig. 11. The dashed line indicates the optimum designs for flat-cover-plate multiweb beams. It is seen that beams with stiffened panels compare well with flat-panel designs in high ranges of loading and show appreciable advantage in low ranges of loading. This result is obtained despite the fact that the criterion for failure of the flat-web design

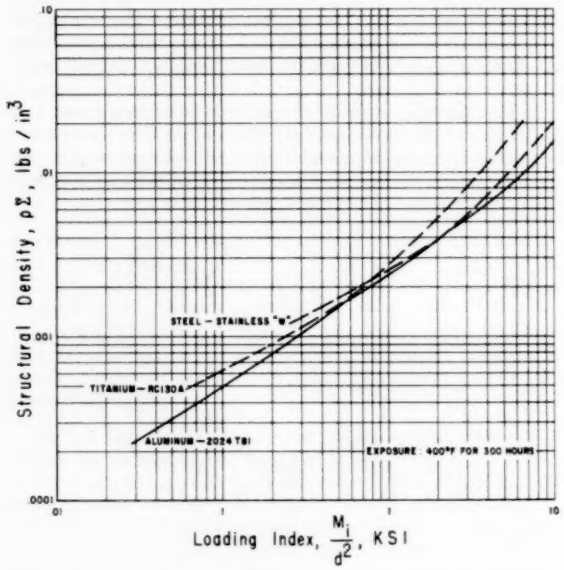


Fig. 8 Comparison of optimum designs for corrugated-web multiweb beams of three different materials

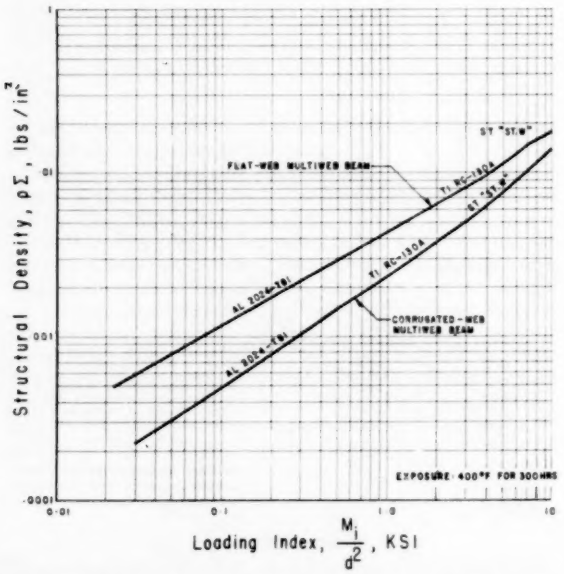


Fig. 9 Comparison of optimum design curves of flat-web and corrugated-web multiweb beams

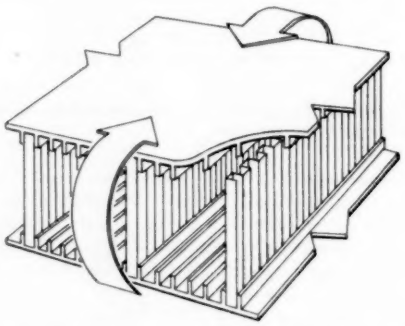


Fig. 10 Integrally stiffened cover panels supported by corrugated webs

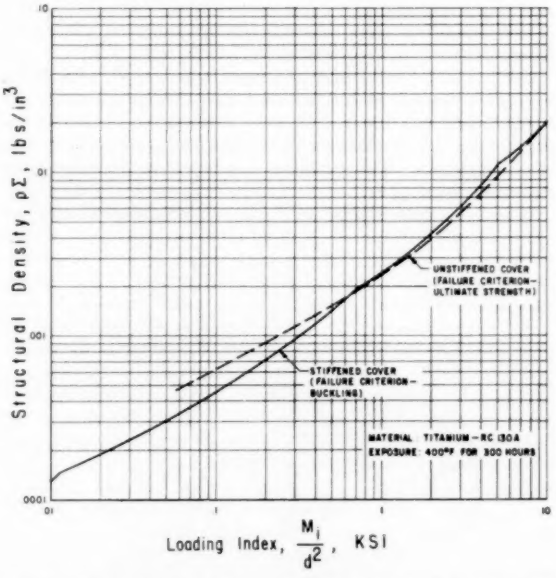


Fig. 11 Comparison of optimum designs of multiweb beams with stiffened covers to those with unstiffened covers

was ultimate panel strength while that for the stiffened panel was instability.

In addition to the advantage of higher efficiency in some ranges of loading, the stiffened panel may offer an advantage in fabrication. If the stiffener is rigidly attached to the cover plate as with currently constructed stiffened plates, the stiffener provides an excellent attachment flange for the webs.

The corrugated-web multiweb beam with sandwich covers illustrated in Fig. 12 was investigated using the sandwich-panel efficiency study of Johnson and Semonian (9) and the web-design procedure previously discussed. The results of this study are shown in Fig. 13.

Two characteristics of the sandwich design are observed:

1 The structural efficiency is extremely high over the entire loading range.

2 Relatively few widely spaced webs may be used without a large weight penalty.

Feasibility of this design rests for the most part on an extensive development program. Results are presented for stainless W steel because problems of fabrication associated with this material are less formidable than for other materials.

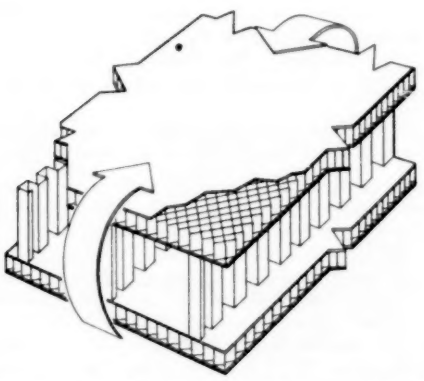


Fig. 12 Sandwich panel supported by corrugated webs

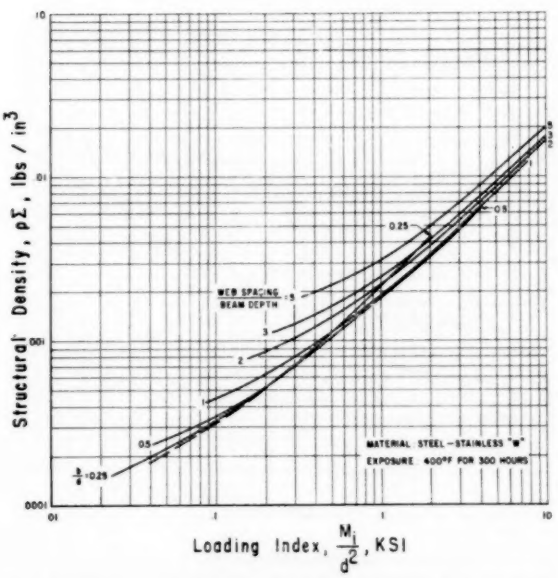


Fig. 13 Efficiency chart for multiweb beam composed of sandwich-panel covers and corrugated supporting webs

The efficiency of the second group of box beams shown in Fig. 1, the wide column-rib supported configurations, was investigated using the general principles discussed earlier. The general instability analysis of Seide and Eppler (10) provided the necessary stiffness criterion for the ribs.

The equations for flexure-induced crushing forces provided the strength criterion for the ribs. Again, corrugated sheet will provide the more efficient compression member.

It is assumed, in the analysis of rib-supported structures, that:

1 The ribs are efficiently joined to the cover panels; i.e., the ribs, stiffness is 100 per cent effective in general instability calculations, and there is no increased weight due to attachments.

2 The chord-to-depth ratio of the rib is sufficiently large to preclude consideration of chord-wise bending stiffness of the rib and other than wide column failure of the compression cover.

3 The tension and compression covers are identical.

In several previous wide-column efficiency analyses (1, 2, 8, 11), buckling strength was expressed as a function of geometry, loading, and material properties as

$$\sigma = F\left(\frac{P_i}{L}\right)^{1/2}(\eta E)^{1/2} \quad \ldots\ldots\ldots\ldots\ldots [3]$$

The factor F represents the efficiency of the geometry of the wide column, P_i/L is the loading index for wide columns, and ηE is the mechanical-properties parameter for the material.

From the present investigation of wide columns with at least one flat surface (e.g., integrally stiffened panels, Z-stiffened panels, sandwich panels, etc.), the practical upper limit of the wide-column efficiency factor, F, appears to be approximately 1.0, but the wide column of highest efficiency does not necessarily lead to the beam of highest efficiency. For instance, it was found that a cover plate with unflanged integral stiffeners having an optimum wide-column efficiency factor of 0.84 is as efficient for beam usage as the Z-stiffened panel having an optimum wide-column efficiency factor of 0.98. The panel with unflanged stiffeners has a greater radius of gyration about the neutral axis of the beam which approximately equalizes the two beam efficiencies.[5]

Using the methods of optimum design outlined and the Z-stiffened wide-column configuration of Farrar (8), efficiency analyses were made of the beam illustrated in Fig. 14. The ef-

[5] A more precise optimization of the stiffened-cover-panel geometry would include the radius of gyration of the box beam cross section as a parameter and would lead to different proportions from those determined by treating the cover panel as a column.

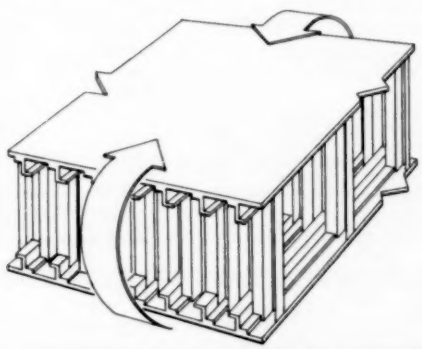


Fig. 14 Z-stiffened panel supported by corrugated ribs

ficiency curves for the configuration are shown in Fig. 15 for RC 130A exposed to 400 F for 300 hr. As in the multiweb designs, as the loading is decreased the support spacing becomes smaller to provide additional stability. In Fig. 16 the envelopes of optimum efficiency are shown for this design in the three example materials of the previous comparisons. Again, the lower density material is the more efficient in the lower ranges of loading while the higher-density materials become the more efficient as loading is increased. It will be shown by comparison that the rib-supported box beam is, ideally, very efficient. It is emphasized, however, that practical and efficient realization of the idealized joining assumed here will be particularly difficult to attain because the ribs are transverse to the plate stiffeners. The sandwich wide column or the plate with unflanged stiffeners will perhaps present fewer attachment problems.

The third group of box beams considered may be described as two facing sheets separated by and efficiently joined to a full-depth core which may be considered a continuous medium.

The example beam shown in Fig. 17 was chosen for an efficiency study. As with the two previous structural categories, three modes of failure are considered:

1 Local buckling of the facing sheet within the boundaries of the support given by the core elements.

2 Local buckling of the elements of the core due to flexure-induced crushing loads.

3 General instability of the composite structure characterized by cylindrical crests and troughs extending over the width of the beam.

Design for the first and second modes of failure is relatively simple; however, no analysis of the cylindrical mode of failure appears in the literature. This mode was therefore analyzed and the strength of the beam in that mode of failure was determined.

Using the principles of optimum design discussed previously an efficiency study was made of this configuration in stainless W

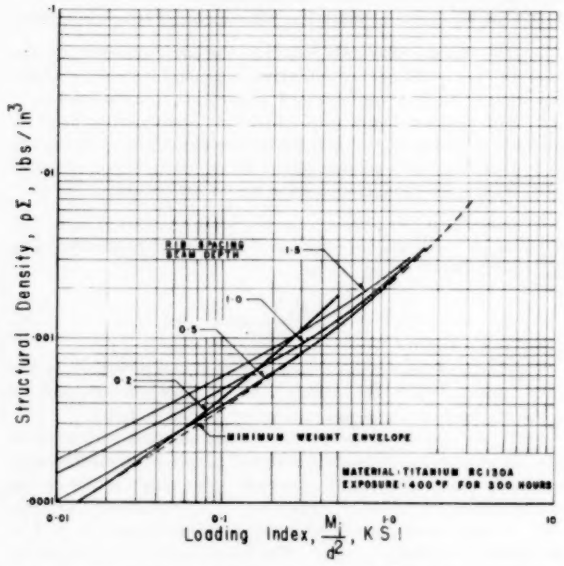


Fig. 15 Efficiency chart for Z-stiffened skin supported on corrugated ribs

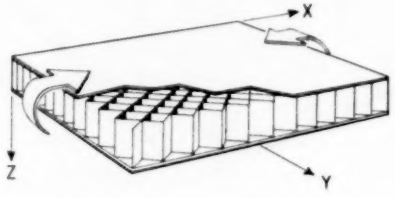


Fig. 17 Full-depth sandwich

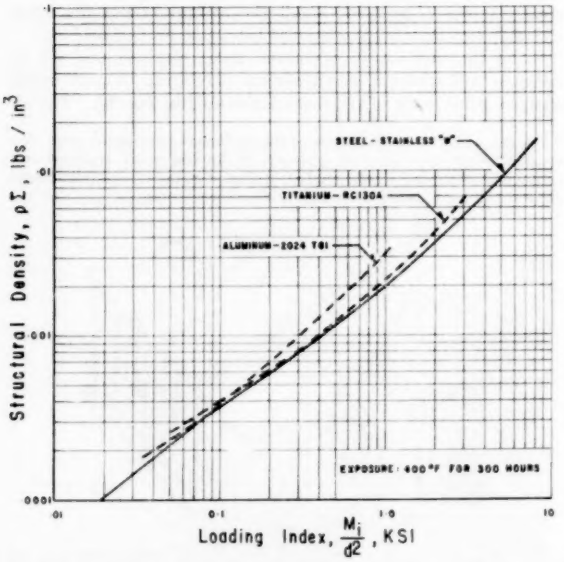


Fig. 16 Comparison of optimum designs for Z-stiffened panels supported by corrugated ribs—beams of three different materials

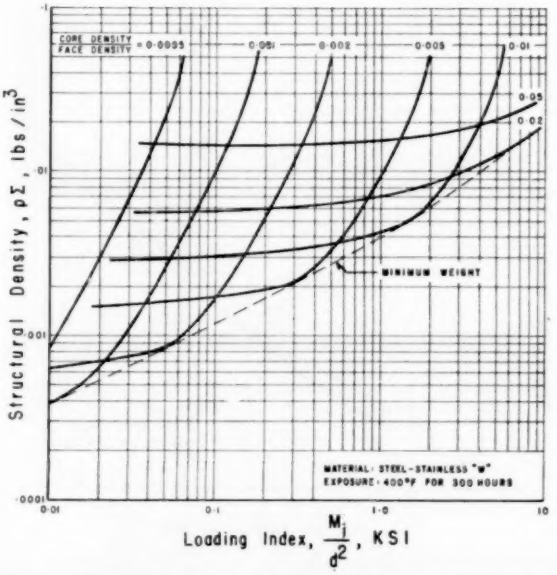


Fig. 18 Efficiency chart for full-depth cellular-core sandwich beam

steel exposed to 400 F for 300 hr. The results of this study are shown in Fig. 18. The parameter of Fig. 18 is core density, but the ratio of cover thickness to beam depth could have been the parameter. The dashed line represents minimum-weight design.

Calculations have shown that the efficiency of the internal medium of the present example may be increased significantly by corrugating the plate elements. This corrugated core has not as yet been incorporated in an analysis of the composite structure, but it is expected that, as in the case of web and rib internal mediums, the increase in core efficiency will reflect in increased beam efficiency.

As with conventional sandwich construction, feasibility of this design rests in the results of an extensive development program. Considerable effort appears to be warranted in this case because of its potential; the configuration can have excellent shear properties in any plane; the covers are continuously supported against deflection; and by use of core that is quasi continuous and of low stiffness in planes parallel to the covers, thermal-stress problems are reduced greatly.

Discussion

A chart is presented in Fig. 19 which shows the optimum efficiency curves for all the configurations treated under the example thermal exposure of 400 F for 300 hr. Similar charts may be computed by the same methods for another given thermal history provided material properties are known for that thermal history.

Except for the sandwich plate on webs, it is seen from Fig. 19 that the stiffened cover plates supported by corrugated ribs is the most efficient design for resisting pure bending. This design has negligible longitudinal shear strength and additional material must be added for that purpose. The same is essentially true for the proposed multiweb designs because those designs are critical or nearly so in crushing, therefore there is no strength available for carrying shear simultaneously with the designed bending load. After adding shear material to the multiweb configurations there is still negligible chord-wise shear stiffness which is not the case with the rib-supported configurations.

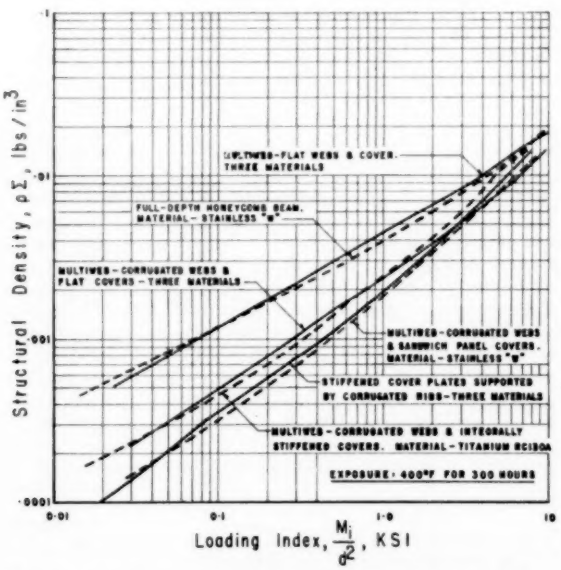


Fig. 19 Comparison of optimum designs for all configurations. Solid lines indicate configurations for which 2024-T81, RC 130 A, and stainless W were considered. Dashed lines indicate configurations for which only one material was considered.

The most obvious disadvantage of the rib-supported structure is the difficulty that will be encountered attaining practically the idealized joining assumed in the analysis. If the cover is a sandwich, the joining problem is no more severe than in the web-supported case. However, when the corrugated rib is to be joined to a cover that is stiffened transversely to the line of the rib joint, serious fabrication problems are encountered. This is especially true for Z-stiffening; however, unflanged integral stiffening appears to present fewer fabrication problems.

The potential efficiency shown by this study indicates that investigation of this attachment problem is warranted.

It is obvious from inspection of the solid lines in Fig. 19 that, for each configuration, optimum efficiency is obtained by going from less dense materials in low loading ranges to more dense materials in high loading ranges.

This trend is explained by the fact that stability is a function of material density. The less dense materials require less supporting structure for stability, therefore they are more efficient in low ranges of loading. In the higher ranges of loading, where the relative weight of the supporting structure is small, materials with high strength-to-density ratios dominate. It is also observed in Fig. 19 that, as the structure becomes more efficient over the entire loading range, the relative weight of the internal structure becomes smaller which permits the material with the higher strength-to-density ratio to dominate over a wider range of loading.

It is concluded, therefore, that the optimum structural material cannot be selected by simply comparing strength-to-density ratios of the available materials. A comparison must be made of the relative efficiencies of the built-up structure in each material.

Similarly, the optimum configuration changes with loading intensity. In low ranges of loading where stability dominates the design, stiffened panels and sandwich plates provide more efficient covers. In high ranges of loading where strength dominates design, the flat cover plate provides the more efficient cover.

The efficiency curve for the full-depth sandwich indicates poor efficiency; however this configuration presents many desirable features which indicate that further investigation toward improving its efficiency may be fruitful.

It is evident from Fig. 19 that further investigation relative to the fabricability of sandwich plates on corrugated webs is certainly warranted on the basis of its very high theoretical efficiency.

The merits of the stiffened panel on corrugated webs as indicated by its high efficiency and potential ease of fabrication indicate that it may find immediate application.

The high structural efficiencies shown by this study involved the assumption that all components were joined efficiently to each other. If attachments are made by welding or other highly efficient joining processes the results are quantitatively valid. Therefore the development necessary to bring high-strength rigid joining into practice is warranted.

Conclusion

Summarizing the results of this investigation:

1 Some new methods are established for the design of aircraft-wing structures.

2 The methods apply either at ambient or elevated temperatures.

3 Corrugated-sheet internal structure is highly desirable because:

(*a*) It is very efficient in compression.

(*b*) It does not lead to significant thermal stresses due to differential expansion of the internal structure and covers.

4 For built-up structures at elevated and room temperatures the choice of the more efficient material is governed not only by comparison of material properties but also by loading index. A comparison of efficiencies must be made in order to determine the most efficient material.

5 The choice of the most efficient configuration depends upon the loading index.

This portion of the efficiency study does not include consideration of the control the designer may have over the structural chord and the associated loading intensity. If there is freedom in choosing the structural chord, the most desirable loading intensity may be selected. An investigation of this aspect of the problem is now being conducted.

Acknowledgment

This research was supported in whole by the United States Air Force under Contract No. AF 33(616)-2810, monitored by the Aircraft Laboratory, Wright Air Development Center.

Bibliography

1 "Effects of Material Distribution on Strength of Panels," by Adam Zahorski, *Journal of the Aeronautical Sciences*, vol. 11, July, 1944, pp. 247–253.

2 "Optimum Number of Webs Required for a Multicell Box Under Bending," by George Gerard, *Journal of the Aeronautical Sciences*, vol. 15, January, 1948, pp. 53–56.

3 "Weight—Strength Analysis of Aircraft Structures," by F. R. Shanley, McGraw-Hill Book Co., Inc., New York, N. Y., 1952.

4 "Charts for Minimum Weight Design of Multiweb Wings in Bending," by E. H. Schuette and J. C. McCulloch, NACA TN 1323, 1947.

5 "Factors Affecting the Design of Thin Wings," by W. J. Conway, Preprint No. 357, SAE Los Angeles Aeronautics Meeting, October 5–9, 1954.

6 "Analysis of the Ultimate Strength and Optimum Proportions of Multiweb Wing Structures," by B. W. Rosen, NACA TN 3633, 1956.

7 "Theory of Elastic Stability," by S. Timoshenko, McGraw-Hill Book Co., Inc., New York, N. Y., 1936, p. 396.

8 "The Design of Compression Structures for Minimum Weight," by D. J. Farrar, *Journal of the Royal Aeronautical Society*, vol. 47, 1943, pp. 1041–1052.

9 "A Study of the Efficiency of High-Strength, Steel, Cellular-Core Sandwich Plates in Compression," by A. E. Johnson and J. W. Semonian, NACA TN 3751, September, 1956.

10 "The Buckling of Parallel Simply Supported Tension and Compression Members Connected by Elastic Deflectional Springs," by Paul Seide and J. F. Eppler, NACA TN 1823, 1949.

11 "The Optimum Design of Compression Surfaces Having Unflanged Integral Stiffeners," by E. J. Catchpole, *Journal of the Royal Aeronautical Society*, vol. 58, 1954, pp. 765–768.

Analysis of the Transient Response of Nonlinear Control Systems

By P. E. W. GRENSTED,[1] CAMBRIDGE, ENGLAND

The calculation by a new analytical method of the transient response of nonlinear control systems is described. If the response is oscillatory, it is possible to obtain expressions for the variation with time of the frequency and amplitude of the oscillation. The response of a servomechanism containing marked saturation, backlash, and coulomb friction has been analyzed successfully.

Nomenclature

The following nomenclature is used in the paper:

A = initial amplitude of oscillations
a = instantaneous amplitude of oscillations
b = semi-backlash zone (Figs. 5 and 6)
c = coefficient of viscous damping
D = operator d/dt
F = mechanical frictional torque (Figs. 5 and 6)
$G(j\omega)$ = transfer function of low-pass linear filter (Fig. 1)
J = moment of inertia of motor shaft (Figs. 5 and 6)
$p(a)$ = describing function of in-phase gain
$q(a)$ = describing function of quadrature gain
T_1, T_2, T_v = time constants of experimental servomechanism (Figs. 5 and 6)
t = time
x = error (Figs. 1, 2, 3, and 4). Motor-shaft rotation, radians (Figs. 5 and 6)
x_0 = initial amplitude of transient
α, β, γ = parameters, defined in Appendix, governing response of experimental servomechanism
μ = instantaneous damping of oscillation
ψ = instantaneous phase of oscillation
ω = instantaneous frequency of oscillation

Introduction

The analysis of nonlinear control systems has proved to be an extremely difficult task. Of the various techniques suggested up to the present time, the "describing-function" method of analysis appears to be one of the most useful. This method is based on the supposition that, if a control system is oscillating in a periodic manner, the signals produced by the nonlinear elements in the system are filtered in the frequency-dependent parts of the system. Thus the harmonics are attenuated in the feedback paths leaving substantially sinusoidal signals at the inputs to the nonlinear elements. The method has been used extensively for the prediction of self-excited oscillations and for the calculation of the frequency and amplitude of these oscillations if they occur (1–11).[2] It also can be used to investigate the response of nonlinear systems to sinusoidal input signals (12, 13), and extensions of the method can be applied to the problem of subharmonic and ultraharmonic resonance (13, 14).

The present paper is concerned to apply a similar method (17) to the calculation of the transient response of nonlinear control systems. The wave form at the input to a nonlinear element is assumed to be a damped oscillation, while the output of the element is considered to be made up of a number of damped oscillations, of which only the one of lowest frequency, equal to the input frequency, is significant.

Differential Equation of a Nonlinear Control System

The method will be described using the system of Fig. 1 as an example. The differential equation of this system will be established first. The forward part of the control loop contains an instantaneous nonlinear element which could represent such phenomena as saturation, dead band, or an on-off (relay or contactor) action. Also hysteresis or backlash may be present in this element. Following this nonlinearity is a low-pass linear filter, with transfer function $G(j\omega)$.

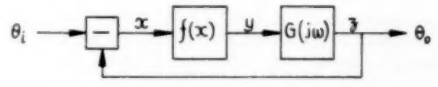


Fig. 1 A Nonlinear Control System

The relation between the input x and the output y of the nonlinear element is

$$y = f(x) \quad [1]$$

while the differential equation relating the input y and output z of the linear filter is

$$S(D)z = R(D)y \quad [2]$$

Here, $S(D)$ and $R(D)$ are polynomials in the differential operator $D \equiv d/dt$ and they are formed from the denominator and numerator, respectively, of the filter-transfer function. Thus

$$R(D)/S(D) \equiv G(D) \quad [3]$$

If the input to the system is zero, the error x and output z of the system are equal and opposite in sign, so that

$$x = -z \quad [4]$$

These equations yield, finally

$$R(D)f(x) + S(D)x = 0 \quad [5]$$

as the differential equation governing the error when the input is zero. This equation also holds for any constant input if the system has zero position lag; i.e., if $G(j\omega)$ has $1/(j\omega)$ as a factor.

The transient behavior of the system in response to a step input is given by the solution of Equation [5] with suitable initial conditions.

Method of Solution

A general solution of Equation [5] is not known. The fol-

[1] I.C.I. Research Fellow, Department of Engineering, University of Cambridge.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-8.

lowing method is essentially approximate, and will be justified by a physical argument.

First, it is supposed that the error will be in the form of a damped oscillation. This is characterized at any instant of time by its amplitude, $a = a(t)$, and phase, $\psi = \psi(t)$. The error is then given by the relation

$$x = a \sin \psi \quad [6]$$

It is also convenient to introduce the concepts of the frequency and damping of this oscillation, even though these quantities may not be constant throughout the transient, by virtue of the nonlinearity in the system. The damping μ is defined as the rate of reduction of amplitude divided by the amplitude; the frequency ω is defined as the rate of change of phase. Then

$$\left.\begin{aligned} \mu &= -\dot{a}/a, \quad \text{or} \quad a = e^{-\int \mu dt} \\ \omega &= \dot{\psi}, \quad \text{or} \quad \psi = \int \omega dt \end{aligned}\right\} \quad [7]$$

An alternative form of Equation [6] is now

$$x = e^{-\int \mu dt} \sin \int \omega dt \quad [8]$$

which may be compared with an oscillation in a linear system

$$x = Ae^{-\mu_0 t} \sin (\omega_0 t + \phi) \quad [9]$$

in which μ_0 and ω_0 are constant. The arbitrary constants A and ϕ in Equation [9] correspond to the arbitrary constants of integration in Equation [8].

As will be shown in examples, the substitution of $a \sin \psi$ for x in Equation [5] enables a relationship between frequency, damping, and amplitude to be established. From this relationship the full transient response can be calculated. But it is necessary first to simplify the term $f(a \sin \psi)$ to one of the form

$$f(a \sin \psi) \approx ap(a) \sin \psi + aq(a) \cos \psi \quad [10]$$

where $p(a)$ and $q(a)$ are functions of amplitude, depending only on the nonlinear element. The justification and significance of this approximation is now discussed.

Justification

An expansion of $f(a \sin \psi)$ by Fourier analysis with respect to phase is[3]

$$\left.\begin{aligned} f(a \sin \psi) &= \sum_{n=1}^{\infty} a_n \sin n\psi + b_0 + \sum_{n=1}^{\infty} b_n \cos n\psi \\ \text{where} \\ a_n &= \frac{1}{\pi} \int_0^{2\pi} f(a \sin \lambda) \sin n\lambda \, d\lambda \\ b_n &= \frac{1}{\pi} \int_0^{2\pi} f(a \sin \lambda) \cos n\lambda d\lambda \\ b_0 &= \frac{1}{2\pi} \int_0^{2\pi} f(a \sin \lambda) \, d\lambda \end{aligned}\right\} \quad [11]$$

The coefficients a_n, b_n, b_0 are functions of a alone.

In this expansion the terms involving a_1 and b_1 represent the "fundamental" component in the output of $f(x)$, b_0 represents a "mean" value, and the remaining terms represent "harmonic" components of greater frequency than the fundamental.

The assumptions in the approximate Equation [10] are first that $f(x)$ is a skew-symmetrical function so that b_0 is zero; also the harmonic components have been ignored. It is supposed that at any instant of time the low-pass filter attenuates the harmonics relative to the fundamental so that they do not modify the wave form significantly at the input to the nonlinear element. This argument is closely analogous to that used to justify a similar approximation, when considering signals of constant frequency and amplitude by means of the describing-function method.

But it should be realized that, while in the transient case the expansion of Equation [11] is still valid, the precise physical significance of the various terms is not certain. In particular, there is no method of establishing exactly the response of the linear filter to the various terms. And so, if ω is varying very rapidly indeed, the relative attenuation of $a_1 \sin \psi$ and $a_3 \sin 3\psi$ may not be as great at a particular instant as that for similar terms with the same constant frequency at that instant. However, provided ω is not varying too rapidly, the relative attenuation is expected to be of the same order as in the constant-frequency case.

For these reasons only the "fundamental" component of $f(a \sin \psi)$ is considered, and by comparing Equations [10] and [11]

$$\left.\begin{aligned} p(a) &= \frac{1}{\pi a} \int_0^{2\pi} f(a \sin \lambda) \sin \lambda \, d\lambda \\ \text{and} \quad q(a) &= \frac{1}{\pi a} \int_0^{2\pi} f(a \sin \lambda) \cos \lambda \, d\lambda \end{aligned}\right\} \quad [12]$$

The gain of the nonlinear element is represented by

$$n(a) = p(a) + jq(a) \quad [13]$$

and is exactly the same as that obtained in an analysis at constant frequency. The quantity $n(a)$ is often termed the describing function of gain, and has been evaluated for most commonly encountered nonlinearities by other authors (15, 16).

Example

As an illustrative example of this method of deriving an equation in terms of frequency, damping, and amplitude, consider a simple velocity-lag controller, with nonlinear gain, and transfer function

$$G(j\omega) = \frac{1}{j\omega(2c + j\omega)}$$

Equation [5] becomes

$$\frac{d^2x}{dt^2} + 2c \frac{dx}{dt} + f(x) = 0 \quad [14]$$

It will be noticed that this is also the equation for the free vibrations of a mass supported on a nonlinear spring, and subject to viscous damping.

On making the substitution $x = a \sin \psi$, and using the approximation of Equation [10], the following equations are derived

$$\omega^2 = p(a) + \mu^2 - 2c\mu - \dot{\mu} \quad [15]$$

$$-\frac{\dot{a}}{a} = \mu = c + \frac{1}{2} \frac{\dot{\omega}}{\omega} + \frac{1}{2} \frac{q(a)}{\omega} \quad [16]$$

The two equations result from equating the coefficients of $a \sin \psi$ and $a \cos \psi$ to zero separately. The $\dot{\mu}$ and $\dot{\omega}$ terms arise as a result of the second differentiation of x with respect to time.

Equation [15] shows that if the system is lightly damped (and the damping is not changing rapidly) the frequency is determined by the instantaneous value of the amplitude

$$\omega^2 \approx p(a) \quad [17]$$

[3] The validity of Equations [11] can be established by regarding a as a constant parameter so that $f(a \sin \psi)$ is periodic in ψ.

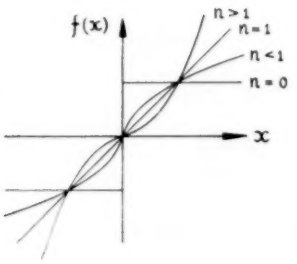


FIG. 2 CHARACTERISTICS OF nTH POWER LAW

Equation [16] shows that the damping can depart from the value c, which would obtain in a linear system whatever the gain. The damping is reduced if $q(a)$ is negative. This occurs when there is a phase lag across the nonlinear element due to backlash or hysteresis. In this respect, Equation [16] expresses a known result in quantitative form. In addition, the damping is modified by the $\dot{\omega}/\omega$ term, being increased if the frequency is increasing during the transient. As will be shown, this effect can contribute to an appreciable part of the damping.

A representative type of nonlinearity which does not introduce phase shift is an nth power law, Fig. 2. In order that $f(x)$ may be a skew-symmetric function for all values of the power, the law will be expressed as

$$f(x) = k^2(\text{sign } x)|x|^n, \quad n \geqslant 0 \ldots\ldots\ldots\ldots [18]$$

The case $n = 0$ corresponds to an ideal on-off element. If $0 < n < 1$, the gain is steadily falling as the amplitude increases giving a "soft-spring" characteristic. A linear characteristic is given by $n = 1$. If $n > 1$ the gain increases with amplitude, giving a hard-spring characteristic.

For this element $q(a) = 0$, while the in-phase gain is

$$p(a) = k^2 C_n a^{n-1} \ldots\ldots\ldots\ldots\ldots\ldots [19]$$

By evaluating the relevant integral in Equations [12] it can be shown (15, 16) that C_n is a constant near unity for a given value of n. Typical values are given in Table 1.

TABLE 1 TYPICAL VALUES OF n AND C_n

n	0	1/2	1	1 1/2	2	2 1/2	3
C_n	$\frac{4}{\pi}$	1.113	1	0.915	$\frac{8}{3\pi}$	0.795	$\frac{3}{4}$

In this case it is possible to find solutions of Equations [15] and [16] when the oscillations are "lightly" damped. From Equation [17]

$$\omega^2 \approx k^2 C_n a^{n-1} \ldots\ldots\ldots\ldots\ldots [20]$$

Using this value of ω in Equation [16] yields

$$-\frac{\dot{a}}{a} = \mu = c + \frac{n-1}{4}\frac{\dot{a}}{a}$$

or

$$\mu = \frac{4}{3+n}c \ldots\ldots\ldots\ldots\ldots [21]$$

So the damping is constant, and the effect of the nonlinear gain is to multiply it by a factor $4/(3 + n)$. For an on-off control ($n = 0$) $\mu = (4/3)c$, and for a cubic hard spring ($n = 3$) $\mu = (2/3)c$. In both these cases a modification of some 30 per cent in the damping is caused by the frequency changing throughout the transient.

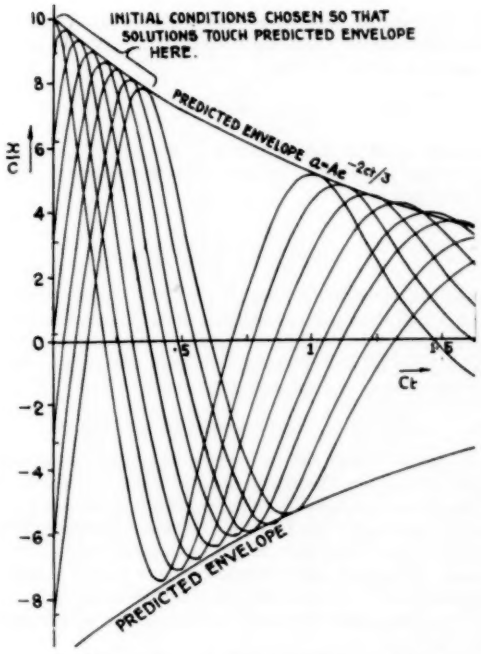


FIG. 3 EXACT SOLUTION OF $\ddot{x} + 2c\dot{x} + x^3 = 0$, WITH PREDICTED ENVELOPE

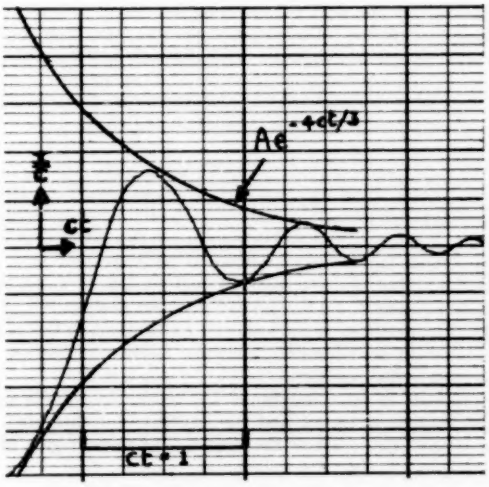


FIG. 4 EXACT SOLUTION OF $\ddot{x} + 2c\dot{x} + (\text{SIGN } x)\ 1 = 0$, WITH PREDICTED ENVELOPE

A physical explanation of this phenomenon can be provided if the system is considered as that of a mass on a nonlinear spring. In the hard-spring case the frequency of oscillations is reduced as the amplitude decreases. Hence, less kinetic energy is required to maintain the oscillations with the result that some is transferred to the potential energy of the spring. So the amplitude decays less rapidly than would be the case if the frequency of oscillations were constant. A similar argument explains the more rapid decay of oscillations with a soft spring.

The iterative solution of Equations [15] and [16] can be re-

peated if greater accuracy is required. Thus a better approximation for the frequency is found from Equation [15] by including the value of μ already derived

$$\omega^2 = k^2 C_n a^{n-1} - 8(n+1)c^2/(3+n)^2 \quad [22]$$

This result could be used to obtain a better value for μ, and so on. But the value of μ obtained so far gives the amplitude of oscillations to be

$$a = Ae^{-\frac{4}{3+n}ct} \quad [23]$$

where A is an arbitrary constant to be settled by the initial conditions. Experiments have shown this to be an adequate solution for the cases $n = 3$ and $n = 0$. In Figs. 3 and 4 exact solutions obtained from a differential analyzer are shown, with the predicted envelope of Equation [23] superimposed.

Application to a Position Control System

A similar procedure can be used for any nonlinear element in a system containing this simple type of transfer function. But only in the case of light damping can the first two steps of the iterative solution of the fundamental equations be considered adequate. The restriction to lightly damped systems is unfortunate in control applications, for here heavy damping is usually required. In order to investigate the accuracy in a typical case, an experimental study was made of a small position-control system.

This servomechanism, shown in Fig. 5, was based on a type T74 Velodyne, which consists of a split-field, d-c fractional-horsepower motor, and a tachometer in one unit. The motor was controlled by its field current supplied from a conventional electronic amplifier. The output signal was taken from a potentiometer geared to the motor shaft through a 100:1 reduction drive. The tachometer output was used for velocity feedback stabilization. A simplified block schematic is shown in Fig. 6. The assumptions on which it is based are described in the following:

The main loop consisted essentially of one major lag T_1, due to the regulation of the armature current in the motor, and an integration between rotational velocity and position. Minor lags were caused by the amplifier and inductance of the motor field windings. These lags were considered to be equivalent to a pure delay, T_2. The velocity feedback was considered to represent a phase advance T_v for purposes of transient analysis.

The nonlinearities in the system were backlash in the gearing of the output potentiometer, nonlinear mechanical friction, and saturation in the amplifier.

Transients of large amplitude (20 revolutions of the motor shaft) were investigated, and the amplifier was saturated, one side or the other most of the time. (The linear region of the amplifier corresponded to $^1/_8$ revolution of the motor shaft.) For this reason, the amplifier was regarded as an on-off element in the analysis.

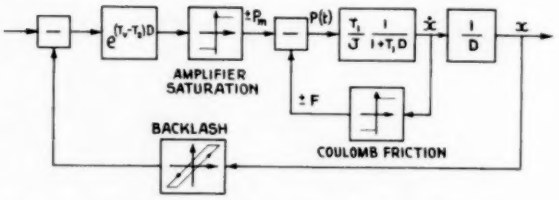


Fig. 6 Simplified Block Schematic of Servomechanism

The mechanical friction was practically independent of speed and represented some 10 per cent of the stand-still torque. It was represented by an on-off element in a negative feedback from speed to torque.

The backlash was eliminated during the experiments as it was not then thought to be possible to include it in the analysis. But provided the backlash zone is small, compared with the transient amplitude, it may be regarded simply as introducing a phase lag inversely proportional to amplitude.

The analysis of the simplified block schematic is given in the Appendix. If the assumption of light damping is made, the envelope of free oscillations of the rotation of the motor shaft is found to be governed by the first-order differential equation

$$-\dot{a} = \frac{2}{3}\left(\frac{1}{T_1}a + \beta a^{1/2} + \gamma - \alpha a^{-1/2}\right) \quad [24]$$

where a is the amplitude of oscillations of the motor shaft in radians. The parameters α, β, and γ result from the minor imperfections of the system, and are defined in the Appendix. Parameter α is directly proportional to the backlash zone, β is directly proportional to the magnitude of mechanical friction, and γ is directly proportional to the time constant of the net phase advance due to velocity feedback less minor lags.

It is shown in the Appendix that time as a function of amplitude may be derived from this equation. But it is instructive to examine the behavior if any one of the three effects dominates. In the following equations A is the initial amplitude of the oscillations.

If backlash and dry friction are ignored, and only minor lags and velocity feedback considered

$$a = (A + T_1\gamma)e^{-2t/(3T_1)} - T_1\gamma \quad [25]$$

giving a finite settling time for the amplitude to become zero.

If dry friction only is considered

$$a = [(A^{1/2} + \beta T_1)e^{-t/(3T_1)} - \beta T_1]^2 \quad [26]$$

also giving a finite settling time.

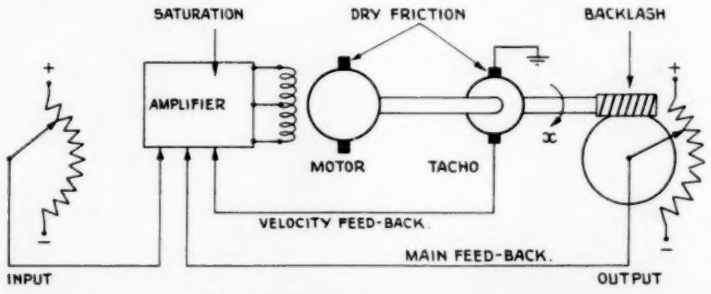


Fig. 5 Experimental Servomechanism

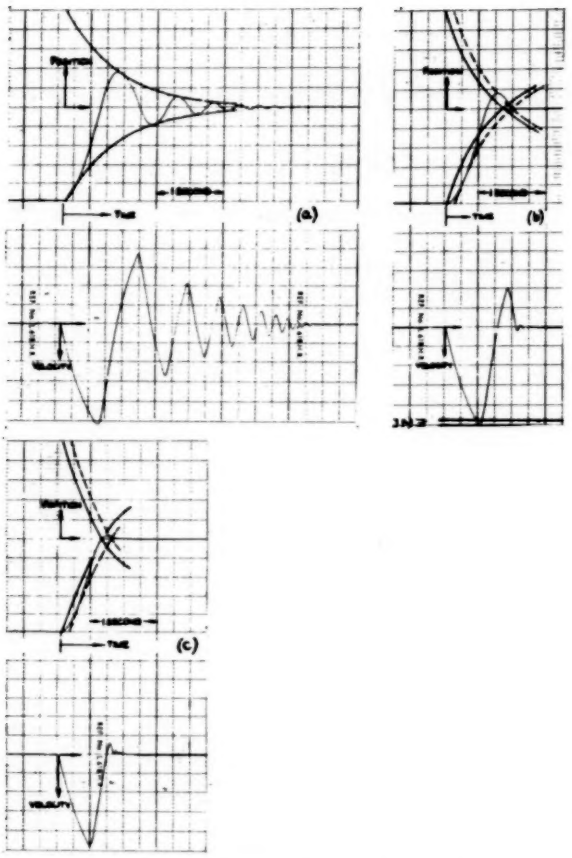


FIG. 7 EXPERIMENTAL RESPONSES WITH PREDICTED ENVELOPES—WITHOUT FLYWHEEL

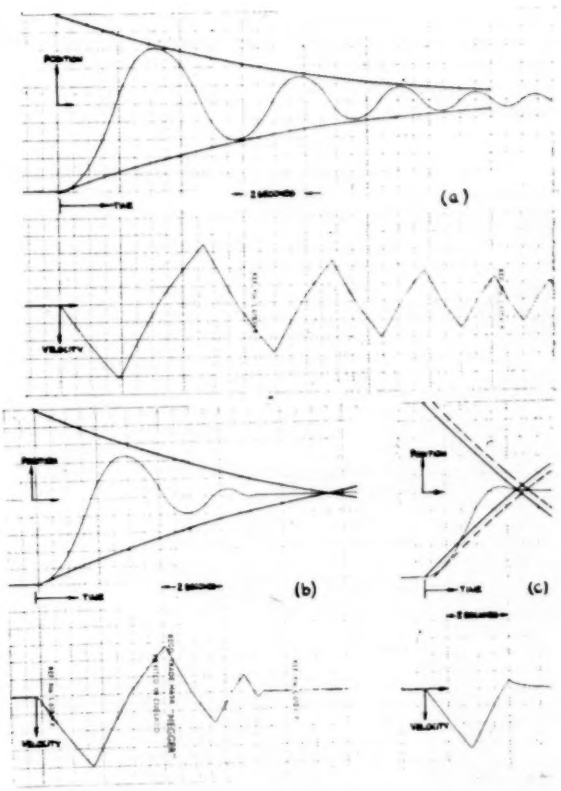


FIG. 8 EXPERIMENTAL RESPONSES WITH PREDICTED ENVELOPES—WITH FLYWHEEL

If backlash only is considered

$$a = [(A^{2/3} - \alpha T_1)e^{-t/T_1} + \alpha T_1]^{3/2} \ldots\ldots\ldots [27]$$

resulting in steady hunting oscillations of amplitude $(\alpha T_1)^{2/3}$ after a long time.

Some results are shown in Figs. 7 and 8, in which predicted envelopes taken from Equations [25]–[27] are superimposed on the experimental responses. The full curves are for an initial amplitude equal to the amplitude x_0 of the step. The broken curves are the same, but shifted in time to allow for the true initial conditions of amplitude and phase, and result in an initial amplitude of oscillation larger than x_0. For the responses of Fig. 8, the inertia of the motor shaft was increased by a factor of 14.7. This resulted in a comparable increase in the major lag T_1 and lower velocities, and so increased the significance of the mechanical friction. The relevant experimental parameters, and equations used, are given in Table 2. In all cases the initial amplitude x_0 was 126 radians.

TABLE 2 PARAMETERS AND EQUATIONS

Fig.	T_1, sec	$T_v - T_2$, sec	γT_1, radians	βT_1, (radians)$^{1/2}$	Equation used
7(*a*)	0.49	0.000	0	1.2	[25]
7(*b*)	0.49	0.073	59	1.2	[25]
7(*c*)	0.49	0.120	97	1.2	[25]
8(*a*)	6.8	0.000	0	4.3	[26]
8(*b*)	6.8	0.107	81	4.3	[25]
8(*c*)	6.8	0.48	361	4.3	[25]

In the heavily damped cases some discrepancy is to be expected because assumptions in the analysis are violated. The discrepancy in Fig. 8(*b*) is attributable to the fact that the damping resulting from minor lags and velocity feedback, and that resulting from dry friction, were of the same order, but the latter contribution has been ignored in computing the envelope.

The results show that settling times are estimated with fair accuracy by this method, even for heavily damped responses. The agreement for lightly damped responses is excellent.

CONCLUSIONS

The restrictions on the type of problem to which this method of analysis may be applied should be mentioned here: (*a*) The assumption that the solution can be regarded as a single damped oscillatory mode implies that the system should be of second order. If additional lags are present they must be regarded as minor ones, introducing phase shift but no attenuation. (*b*) Only in lightly damped systems can a frequency-amplitude relation be established immediately. This is a necessary first step in determining the amplitude as a function of time. The damping and hence the amplitude must depend on the rate of change of frequency if a marked nonlinearity is present in the main loop of the system.

The examples of this paper have shown that, in suitable cases,

the full transient response of nonlinear systems can be established in an approximate analytical form. Moreover, parameters of the problem are retained in the solution. Hence this method can assist in the synthesis as well as the analysis of nonlinear control systems.

BIBLIOGRAPHY

1 "A Method of Analyzing the Effect of Certain Kinds of Nonlinearity in Closed-Cycle Control Systems," by A. Tustin, *Journal of the Institution of Electrical Engineers*, vol. 94, part IIA, 1947, pp. 152–160.

2 "On Some Nonlinear Phenomena in Regulating Systems," by L. C. Goldfarb, *Avtomatika i Telemekhanika*, vol. 8, 1947, pp. 349–383.

3 "A Frequency Response Method of Analyzing and Synthesizing Contactor Servomechanisms," by R. J. Kochenburger, Trans. AIEE, vol. 69, part I, 1950, pp. 270–283.

4 "Sinusoidal Analysis of Feedback-Control Systems Containing Nonlinear Elements," by E. C. Johnson, Trans. AIEE, vol. 71, part II (Applications and Industry), 1952, pp. 169–181.

5 "Some Saturation Phenomena in Servomechanisms," by E. Levinson, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 1–9.

6 "Limiting in Feedback-Control Systems," by R. J. Kochenburger, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 180–192.

7 "Coulomb Friction in Feedback-Control Systems," by V. B. Hass, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 119–123.

8 "Open-Loop Frequency Response Method for Nonlinear Servomechanisms," by R. L. Cosgriff, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 222–225.

9 "Backlash in a Velocity Lag Servomechanism," by N. B. Nichols, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 462–467.

10 "Approximate Frequency-Response Methods for Representing Saturation and Dead Band," by H. Chestnut, Trans. ASME, vol. 76, 1954, pp. 1345–1363.

11 "Stability Characteristics of Closed-Loop Systems With Dead Band," by C. H. Thomas, Trans. ASME, vol. 76, 1954, pp. 1365–1382.

12 "The Frequency Response of a Certain Class of Nonlinear Feedback Systems," by J. C. West and J. L. Douce, *British Journal of Applied Physics*, vol. 5, 1954, pp. 204–209.

13 "The Dual Input Describing Function and Its Use in the Analysis of Nonlinear Feedback Systems," by J. C. West, J. L. Douce, and R. K. Livesley, Proceedings of the Institution of Electrical Engineers, vol. 103, part B, 1956, pp. 463–473.

14 "The Mechanism of Subharmonic Generation in a Feedback System," by J. C. West and J. L. Douce, Proceedings of the Institution of Electrical Engineers, vol. 102, part B, 1955, pp. 569–574.

15 "Describing Function Method of Servomechanism Analysis," by H. D. Greif, Trans. AIEE, vol. 72, part II (Applications and Industry), 1953, pp. 243–248.

16 "On a Method for Investigating Nonlinear Oscillations and Control Systems," by K. Magnus, *VDI-Forschungsheft*, series B, vol. 21, 1955, pp. 451–483.

17 "The Frequency Response Analysis of Non-Linear Systems," by P. E. W. Grensted, Proceedings of the Institution of Electrical Engineers, vol. 102, part C, 1955, pp. 244–255.

Appendix

The analysis of the simplified block schematic of Fig. 6 is given in the following. In the diagram, x is the motor shaft rotation in radians, J is the moment of inertia of the rotating parts, P_m is the gross stand-still torque, F is the constant mechanical frictional torque. The time constant T_1 results from the armature-current regulation. The time constant $T_v - T_2$ is that of the net phase advance resulting from the velocity feedback, T_v and minor lags T_2. The semi-backlash zone in the gearing to the output potentiometer is b radians rotation of the motor shaft. Let

$$x = a \sin \psi$$

then

$$\dot{x} = \dot{a} \sin \psi + a\omega \cos \psi$$

and

$$\ddot{x} = (\ddot{a} - a\omega^2) \sin \psi + (2\dot{a}\omega + a\dot{\omega}) \cos \psi$$

By working backwards from $\dot{x}$ in the block schematic, the instantaneous value of the generated torque less coulomb friction is

$$P(t) = \frac{J}{T_1}(\dot{x} + T_1\ddot{x}) = \frac{J}{T_1}\{[\dot{a} + T_1(\ddot{a} - a\omega^2)] \sin \psi + [a\omega + T_1(2\dot{a}\omega + a\dot{\omega})] \cos \psi\} \quad \ldots\ldots [28]$$

By working forward from x around the main feedback loop, and noting that the phase advance due to the velocity feedback, minor delays, and backlash is $[(T_v - T_2)\omega - b/a]$ if $a \gg b$ and $(T_v - T_2) \ll 1/\omega$

$$P(t) = -\frac{4}{\pi} P_m \sin [\psi + (T_v - T_2)\omega - b/a] - \frac{4}{\pi} F \sin [\text{phase of } \dot{x}]$$

or

$$P(t) = -\frac{4}{\pi} P_m \sin \psi - \frac{4}{\pi} P_m[(T_v - T_2)\omega - b/a] \cos \psi - \frac{4}{\pi} F \frac{\dot{a} \sin \psi + a\omega \cos \psi}{(\dot{a}^2 + a^2\omega^2)^{1/2}} \quad \ldots\ldots [29]$$

to this order of accuracy.

Equating Equations [28] and [29] and neglecting $(\dot{a}/a)^2$ in comparison with ω^2, yields

$$\omega^2 = \frac{4}{\pi} \frac{P_m T_1}{J} \frac{1}{a} \quad \ldots\ldots [30]$$

and

$$-\frac{\dot{a}}{a} = \frac{1}{2T_1} + \frac{2}{\pi} \frac{F}{J} \frac{1}{\omega a} + \frac{1}{2} \frac{\dot{\omega}}{\omega} + \frac{2}{\pi} \frac{P_m}{J} \frac{(T_v - T_2)}{a} - \frac{2}{\pi} \frac{P_m}{J} \frac{b}{a^2\omega^2} \quad \ldots\ldots [31]$$

Thus, if the damping is light, the frequency is inversely proportional to the square root of amplitude.

Using Equation [30] in Equation [31] yields

$$-\dot{a} = \frac{2}{3}\left(\frac{1}{T_1} a + \beta a^{1/2} + \gamma - \alpha a^{-1/2}\right) \quad \ldots\ldots [32]$$

where

$$\alpha = b\left(\frac{4P_m}{\pi J}\right)^{1/2}; \quad \beta = F\left(\frac{4}{\pi J P_m}\right)^{1/2}; \quad \gamma = (T_v - T_2)\frac{4P_m}{\pi J}$$

In the text, the solution of this equation for the envelope of transient oscillations is given when any two of the parameters α, β, or γ are zero. The exact solution of Equation [32] (solved by separation of variables) expresses time as a function of amplitude

$$t = -3T_1 \sum \frac{\lambda_{1,2,3}^2}{(\lambda_{2,3,1} - \lambda_{1,2,3})(\lambda_{3,1,2} - \lambda_{1,2,3})} \log\left(\frac{a^{1/2} - \lambda_{1,2,3}}{A^{1/2} - \lambda_{1,2,3}}\right)$$

The summation consists of three terms in which the first, second, or third suffix is used, respectively. A is the initial amplitude when $t = 0$, and λ_1, λ_2, λ_3 are the (unequal) roots of

$$\xi^3 + T_1\beta\xi^2 + T_1\gamma\xi - T_1\alpha = 0$$

Algebraic Approach to Design of Automatic Controls

By RUFUS OLDENBURGER,[1] LAFAYETTE, IND.

The solution of linear problems in automatic control is generally reduced to the study of ordinary differential equations and thus to characteristic equations, which are algebraic. It is then necessary to solve an algebraic equation to predict the transients for a given controlled system. Whether or not the system is stable can be determined from the simple test of Routh. Because of the great difficulties encountered in the past in solving algebraic equations, especially in the case of all roots complex, resort has been made to qualitative methods of automatic control design based on frequency response and other techniques. The author applies to algebraic equations for stable systems certain procedures, including a right to left synthetic division, which enable the engineer to approximate some of the roots, after which the solution for all of the roots can be obtained readily. From the roots, the analyst can tell what the transients will look like. Good transients are necessary for good control. The author's method, in use at the Woodward Governor Company for several years, is applied to the design of a governor for a gas turbine.

1 Introduction

PHYSICAL problems often reduce to an ordinary linear differential equation with real coefficients and hence to the solution of a characteristic equation which is algebraic. In the case of a linear system with a variable under automatic control, an ordinary differential equation generally relates the controlled variable to the setting of this variable and such an equation also relates the controlled variable to a disturbing quantity. Thus in the control of engine rpm an ordinary differential equation can be written to give rpm in terms of the speed setting, and another to connect rpm to load.

The lack of efficient methods of solving algebraic equations has led to qualitative indirect design rules for automatic controls, prominent among which are those of frequency response (1).[2] It also has led to the examination of roots by the use of curves (2) the drawing of which is necessarily time-consuming.

The major difficulty in working with algebraic equations lies in solving equations with all roots complex. Graeffe's (3) and other classical procedures are quite tedious. The methods of Lin (8), Lyon, Ku, Woodruff, Hitchcock, and Koenig require a formal cut-and-try procedure which leaves little room for the exercise of judgment. The techniques used here have been standard at the author's former company for several years, and have been found to reduce control design time substantially.

A control system is designed to be stable. This means that the real parts of the roots of the corresponding characteristic equation are negative. Such an equation the writer calls *stable*. There is a simple test due to Routh (4) to determine whether or not a system and its equation are stable. The author's method of solving equations depends on the discovery of certain algebraic properties of stable equations. The actual values of the roots are established by synthetic division. A brief note of some aspects of the method appeared in the author's discussion of a paper by W. R. Evans (5).

The automatic-control scientist is generally concerned with stable systems. If, however, he has an unstable system he can still use the author's method of solution provided he first reduces the given equation to one that is stable. This is done by a well-known division process.

From the roots the control expert can tell rather precisely what the transient will look like when his system is disturbed. Certain roots will always be dominant. These are the roots with the numerically smallest real parts.

2 Origin of Equations

Consider a physical system with an input $m(t)$ and output $c(t)$. If the system is linear the relation between m and c is normally (6) given by the differential equation

$$a_0 \frac{d^n c}{dt^n} + a_1 \frac{d^{n-1} c}{dt^{n-1}} + \ldots + a_{n-1} \frac{dc}{dt} + a_n c = b_0 \frac{d^r m}{dt^r} + b_1 \frac{d^{r-1} m}{dt^{r-1}} + \ldots + b_{r-1} \frac{dm}{dt} + b_r m \quad [1]$$

for real numbers $a_0, a_1, \ldots, a_n$, and $b_0, b_1, \ldots, b_r$. The values of c and m can be taken as deviations from a steady state for which $c = m = 0$. If m is suddenly changed from a nonzero value to $m = 0$ the equation

$$a_0 \frac{d^n c}{dt^n} + a_1 \frac{d^{n-1} c}{dt^{n-1}} + \ldots + a_{n-1} \frac{dc}{dt} + a_n c = 0 \quad [2]$$

holds, the solution of which is

$$c = C_1 e^{\alpha_1 t} + C_2 e^{\alpha_2 t} + \ldots + C_n e^{\alpha_n t} \quad [3]$$

where the α's are the roots of the algebraic equation

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0 \quad [4]$$

The terms corresponding to a pair of roots $-a \pm bj$ for $j = \sqrt{-1}$ and $a, b > 0$ can be grouped into $e^{-at}[A \cos bt + B \sin bt]$ for constants A and B. The imaginary part b is $2\pi f$ for the frequency f of the corresponding oscillation. The C's depend on the initial values of c and its first $(n - 1)$ derivatives; that is, on their values at the time of the disturbance. In what follows it is assumed that $a_0 > 0$.

Such a system is said to be *stable* if, for every "disturbance" m that dies out, the response of the output c also dies out. This means that the characteristic roots $\alpha_1, \ldots, \alpha_n$ must have negative real parts. To determine whether or not the given system is stable we apply Routh's theory as follows: We form the array (shown for n even)

---

[1] Professor of Electrical and Mechanical Engineering, Purdue University. Mem. ASME.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–9, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, December 27, 1956. Paper No. 57—IRD-2.

$$a_0 \quad a_2 \quad a_4 \ldots a_{n-2} \quad a_n$$

$$a_1 \quad a_3 \quad a_5 \ldots a_{n-1}$$

From the two rows we form a third row as shown

$$a_0 \qquad a_2 \qquad a_4 \qquad a_6 \ldots a_{n-2} \quad a_n$$

$$a_1 \qquad a_3 \qquad a_5 \qquad a_7 \ldots a_{n-1}$$

$$\left(a_2 - \frac{a_0 a_3}{a_1}\right) \quad \left(a_4 - \frac{a_0 a_5}{a_1}\right) \quad \left(a_6 - \frac{a_0 a_7}{a_1}\right) \quad \ldots a_n$$

Where an entry is missing it is understood to be zero. Thus there is no entry in the second row and last column (under a_n). The quantity in the third row under a_{n-1} is thus

$$\left(a_n - \frac{a_0 \cdot 0}{a_1}\right)$$

or a_n. A fourth row is formed from the second and third rows in the same way as the third row is formed from the first and second. Similarly, other rows are formed from the two preceding ones. Continuing this process we thus form a Routh array.

For a cubic the Routh array is

$$a_0 \qquad a_2$$

$$a_1 \qquad a_3$$

$$\left(a_2 - \frac{a_0 a_3}{a_1}\right)$$

and for a quartic

$$a_0 \qquad a_2 \qquad a_4$$

$$a_1 \qquad a_3$$

$$\left(a_2 - \frac{a_0 a_3}{a_1}\right) \qquad a_4$$

$$\left[a_3 - \frac{a_1 a_4}{\left(a_2 - \frac{a_0 a_3}{a_1}\right)}\right]$$

For the system to be stable, it is well known that the coefficients of Equation [4] must be positive. This is assumed throughout the paper. According to Routh the system and the corresponding Equation [4] are then stable if and only if all the entries in the first column of the Routh array are positive.

Unless stated otherwise, *it will be assumed in what follows that the equations are stable.*

It is always possible to take $a_0 = 1$ in Equation [4], whence we have

$$x^n + a_1 x^{n-1} + \ldots + a_{n-1} x + a_n = 0 \ldots\ldots [5]$$

In what follows, our equations will be taken in this form. Thus for a cubic we have

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0 \ldots\ldots [6]$$

3 Real Roots

The solution of a stable algebraic equation for real roots is quite simple (7). Since the sum of the roots of Equation [5] is $-a_1$ and no roots are positive, it follows that the real roots are between 0 and $-a_1$. We may consider the solution of the first two terms of Equation [5] set equal to zero

$$x + a_1 = 0 \ldots\ldots [7]$$

as an approximation to the numerically largest real root (or roots) of Equation [5]. Since part of a_1 must be employed for the other roots, it is generally advisable to compensate and use something like $-a_1/2$ in place of $-a_1$. The numerically smallest root (or roots) may be approximated by the equation

$$a_{n-1} x + a_n = 0$$

obtained by setting the last two terms of Equation [5] equal to zero. This approximation is thus $-a_n/a_{n-1}$. For the proof the reader is referred to reference (7).

Real roots are removed by synthetic division. It is generally convenient to solve for the numerically largest real roots first, using the author's right to left synthetic division, illustrated in the example to follow.

We shall solve

$$x^3 + 13x^2 + 39x + 27 = 0 \ldots\ldots [8]$$

From the first two terms (see Equation [7]) we obtain the trial root -13. We try the more convenient number -10 instead. The division follows

1	13	39	27	|10
	9.37	36.3	27	
	3.63	2.7	0	

Here we first divided the trial divisor 10 into 27 to obtain the remainder 2.7. Subtracting 2.7 from the coefficient 39 we obtain the difference 36.3. Dividing 10 into 36.3 we obtain the remainder 3.63. Subtracting this from the entry 13 we obtain the difference 9.37. If the division were exact this difference would be equal to the trial divisor 10. It is not, and we can take 9.37 as the next trial divisor. Taking a little less, conveniently 9, we have the division

1	13	39	27	|9
	9	36	27	
	4	3	0	

The division is now exact, whence -9 is a root of Equation [8]. The remainder zero and the entry 27 above it were included in the foregoing divisions for the purpose of exposition, and are normally not written. The remaining roots are roots of the reduced equation

$$x^2 + 4x + 3 = 0$$

formed from the remainders of the division.

We could have solved first for the numerically smallest real root as follows: The last two terms in Equation [8] give the trial root $-27/39$, or -0.69. We increase this a little numerically and try -0.8 as follows, using ordinary (left-to-right) synthetic division

1	13	39	27	|0.8
	0.8	9.76	23.4	
1	12.2	29.2	3.6	

Here we have rounded off the numbers to three digits. The presence of a remainder 3.6 shows that 0.8 should be increased numerically. Trying the root -1 we have

1	13	39	27	|1
	1	12	27	
1	12	27		

whence $x = -1$ is a root of Equation [8].

An odd-degree equation always has at least one real root. The real roots of an equation should be removed before solving for the complex roots. A few trials will indicate whether or not real roots remain.

4 A Problem From Industry

For higher-degree equations there is a good chance that the root with the numerically largest real part is real. In fact, in the design of governors for prime movers a typical set of roots (rounded off) is given by the following

$$-60,\ -30,\ -4,\ -0.9,\ -1 \pm j \ldots\ldots\ldots\ldots [9]$$

The equation which gave these roots was

$$x^6 + 103x^5 + 2800x^4 + 18{,}000x^3 + 41{,}000x^2 + 46{,}000x + 20{,}000 = 0 \ldots\ldots [10]$$

This equation can be solved as follows: The first two terms yield the trial root

$$x = -103$$

We compensate and take -50 instead. The division could be performed by a rigid left-to-right solution procedure, but intelligent cut and try using approximate figures in the right-to-left technique is better. The author originally solved Equation [10] for the numerically largest root as follows.

1	103	2800	18,000	41,000	46,000	20,000	|50
	53	2500	17,000	40,000	45,600		
	50	340	800	910	400		
	61						|60
1	42	280	670	760	330		

Here the first entry 53 of the second row was overcompensated to yield 60 as the second trial divisor. Unnecessary numbers are not repeated.

If the first entry in the second row of a right-to-left division is definitely greater than the divisor for this division, one should make it still greater to obtain the next trial divisor, and if definitely less, then still less. A rule for the amount of compensation has not been determined.

From the remainders of the last division we obtain the reduced equation

$$x^5 + 42x^4 + 280x^3 + 670x^2 + 760x + 330 = 0 \ldots [11]$$

The first two terms yield the trial root

$$x = -42$$

Compensating, the divisor 20 is employed instead. The division (right-to-left) follows

1	42	280	670	760	330	|20
	30	250	630	740		
	12	32	37	16		
	33	260	640	750		|30
	9	21	25	11		
	34	260	650	750		|34
1	7.6	19	22	9.7		

The entry 30 suggests the second trial divisor 30 and the entry 33 suggests the third trial divisor 34. Clearly, -34 is a root of Equation [11], and hence of Equation [10]. The reduced equation is now

$$x^4 + 7.6x^3 + 19x^2 + 22x + 9.7 = 0 \ldots\ldots [12]$$

The roots of Equation [12] are found in a similar manner.

The roots of Equation [10] are thus given by the array [9] to one significant digit. If Equation [10] is associated with a differential Equation [2] the solution of the latter equation is

$$c = C_1e^{-60t} + C_2e^{-30t} + C_3e^{-4t} + Ce^{-0.9t} + e^{-t}[A \cos t + B \sin t] \ldots\ldots [13]$$

The constants C_1, C_2, C_3, C, A, and B depend on the initial conditions, say at $t = 0$; that is, on the values of c, dc/dt

$$d^2c/dt^2,\quad d^3c/dt^3,\quad d^4c/dt^4 \text{ and } d^5c/dt^5 \text{ at } t = 0$$

As t increases the terms in e^{-60t}, e^{-30t}, and e^{-4t} die out very rapidly compared to the remaining terms in Equation [13], and therefore they can be dropped. The solution is now

$$c = Ce^{-0.9t} + e^{-t}(A \cos t + B \sin t)$$

Let time be in seconds. It can be seen that the transient will die out, for practical purposes, in about 5 sec, and that the solution will be oscillatory, but not much, with a frequency of about 0.2 cps.

5 Complex Roots

For a stable quartic[3]

$$x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0 \ldots\ldots [14]$$

with all complex roots *one of the quadratics* $(x^2 + a_1x + a_2)$, $(a_3x^2 + a_3x + a_4)$ *is approximately a factor of the left member of Equation* [14]. The first quadratic is a factor for which the sum of the corresponding pair of roots is a maximum in absolute value. In fact Equation [14] always has a factor $(x^2 + \alpha x + \beta)$, where

$$\frac{a_1}{2} \leqq \alpha < a_1,\quad \frac{a_2}{6} \leqq \beta < a_2 \ldots\ldots [15]$$

or

$$\frac{a_1}{2} \leqq \alpha < a_1,\quad \beta < \frac{a_2}{6},$$

$$\frac{a_4}{a_2} < \beta \leqq \frac{6a_4}{a_2},\quad \frac{a_3}{6a_4} \leqq \frac{\alpha}{\beta} < \frac{a_3}{a_4}$$

Theory and experience show that it is generally advisable to use α about midway between $a_1/2$ and a_1, and β about midway between a_2 and $a_2/6$, say $a_2/3$, with right-to-left division.

The italicized statement above is similar to one given by von Karman and Biot (9), but these authors assume that the absolute values of two of the roots dominate those of the remaining pair. We shall solve the equation

$$x^4 + 3x^3 + 5x^2 + 4x + 2 = 0 \ldots\ldots [16]$$

Routh's criterion shows that this equation is stable.

From the first three terms of Equation [16], using the first row of Inequalities [15] we obtain the trial divisor

$$x^2 + 2x + 3$$

We use the author's right-to-left division process as follows

1	3	5	4	2	|1, 2, 3
		0.67	1.3	2	|1, 0.9, 0.67
		4.33	2.7	0	
	0.9	1.8	2.7		
	2.1	2.53	0		
1	2	3			
	0.1	−0.47			

The coefficients 1, 2, 3 of the divisor are written at the right.

[3] For the proof see the Appendix.

The quotient 0.67 is obtained by dividing the last divisor coefficient 3 into the last coefficient 2 of the given quartic. Multiplying this quotient 0.67 by the divisor 1, 2, 3 yields the entries 0.67, 1.3, 2 written in the second row. Subtracting these entries from the numbers above them we obtain the remainders 4.33, 2.7, 0 in the third row. The zero may be omitted. Dividing 3 of the divisor into the remainder 2.7 we obtain the quotient 0.9, written before the 0.67 on the right. Multiplying the quotient 0.9 by the divisor yields the entries 0.9, 1.8, 2.7 of the fourth row. Subtracting from the numbers above them we obtain the remainders 2.1, 2.53, 0. The first coefficient of the quotient should of course be 1. We therefore write 1 as the leading coefficient of the quotient. Multiplying this by the divisor yields the entries 1, 2, 3 of the sixth row. Subtracting from the numbers above gives the remainders 0.1 and -0.47 as shown. In practice the last two rows of the division are not written.

If the division were exact we should have the remainders 2 and 3 in place of 2.1 and 2.53 in the fifth row and zero remainders in the seventh row. From these fifth-row remainders we obtain

$$x^2 + 2.1x + 2.53$$

which suggests our next trial.

We compensate and use

$$x^2 + 2x + 2$$

instead. We have

1	3	5	4	2	1, 2, 2
		1	2	2	1, 1, 1
		4	2		
	1	2	2		
	2	2			

The division is now exact, whence the divisor

$$(x^2 + 2x + 2)$$

and the quotient

$$(x^2 + x + 1)$$

are the factors of the left side of Equation [16].

6 High-Degree Equations

We consider the sextic equation

$$x^6 + a_1x^5 + a_2x^4 + a_3x^3 + a_4x^2 + a_5x + a_6 = 0 \ldots [17]$$

We assume this to be stable with all complex roots. As in the case of quartics, it can be shown that *at least one of the quadratics*

$$x^2 + a_1x + a_2, \quad a_2x^2 + a_3x + a_4, \quad a_4x^2 + a_5x + a_6$$

is a fair approximation to a quadratic factor of the left side of Equation [17] (for which the sum of the corresponding roots is a maximum in absolute value).

To illustrate the solution of high-degree equations we consider the sextic

$$x^6 + 6x^5 + 18x^4 + 31x^3 + 34x^2 + 22x + 8 = 0 \ldots [18]$$

This is a stable equation with all roots complex. The first three terms yield the trial divisor

$$x^2 + 6x + 18$$

We take convenient coefficients a little less and use

$$x^2 + 5x + 10$$

instead. We have

1	6	18	31	34	22	8	1,	5, 10
				0.8	4		1, 1.71,	2.42, 1.8, 0.8
				33.2	18			
			1.8	9				
			29.2	24.2				
		2.42	12.1					
		15.58	17.1					
	1.71	8.55						
	4.29	7.03						

For the division to have been exact we should have had 5, 10 in place of the remainders 4.29, 7.03. The remainders suggest the divisor

$$x^2 + 4.29x + 7.03$$

We overcompensate and try instead

$$x^2 + 4x + 6$$

as follows

1	6	18	31	34	22	8	1,	4, 6
				1.3	5.2		1, 2.2,	3.7, 2.8, 1.3
				33	17			
			2.8	11				
			28	22				
		3.7	15					
		14	13					
	2.2	8.8						
	3.8	5.2						

Here we have rounded off the numbers to two places. The remainders suggest the divisor

$$x^2 + 3.8x + 5.2$$

Compensating the coefficients we try

$$x^2 + 3x + 5$$

instead. The division follows

1	6	18	31	34	22	8	1,	3, 5
				1.6	4.8		1, 3,	4.4, 3.4, 1.6
				32	17			
			3.4	10				
			28	22				
		4.4	13					
		14	15					
	3	9						
	3	5						

The remainders are identical with the divisor. Thus

$$x^2 + 3x + 5$$

is a factor of the left side of Equation [18]. The quotient

$$x^4 + 3x^3 + 4.4x^2 + 3.4x + 1.6$$

is the other factor.

Had we used the last three terms $34x^2 + 22x + 8$ of the left side of Equation [18] as a trial factor, we would have employed left-to-right synthetic division as in the method of Lin. Had we tried the middle terms $18x^4 + 31x^3 + 34x^2$ we would have used right-to-left division.

7 Comments on Solutions

In solving for complex roots by using the leading three terms of

Equation [5] for a trial divisor, we decrease the coefficients first, because if these terms approximate an actual factor, the coefficients of this factor are less than a_1 and a_2 respectively. Similarly, if we wish to use

$$x^2 + \frac{a_{n-1}}{a_{n-2}}x + \frac{a_n}{a_{n-2}}$$

for a trial divisor, we first increase the coefficients. Theoretical considerations can be given to justify the manner in which remainders after divisions in the examples were compensated to yield better succeeding trial divisors. These considerations are complicated and will be omitted.

The theory of this paper can be generalized to higher-degree equations. However, the approximations obtained by using three successive terms of the left side of Equation [5] become poorer and poorer as we go to the eighth, tenth, and higher degrees. Also, in the author's experience with numerous equations from all major fields of engineering, he has never encountered equations of degree higher than the sixth with all-complex roots. That this should be so can be justified on the basis of statistical theory (7). We therefore omit the solution of eighth and higher-degree equations with all-complex roots.

Every seventh-degree equation with real coefficients has at least one real root. When this is removed one obtains a sixth-degree equation, for which the solution has been described in detail in this paper.

If in Equation [5] a_1 is very large relative to 1 and not small compared to the other coefficients, the term x^n can be dropped from the equation, and the equation reduced to one of lower degree. Thus

$$x^4 + 100x^3 + 200x^2 + 200x + 10 = 0$$

can be reduced immediately to

$$100x^3 + 200x^2 + 200x + 10 = 0$$

One root is actually near -100. Similarly if a_n is small compared to the other coefficients we may drop a_n. Thus we may omit the term 10 in the last equation to obtain

$$100x^2 + 200x + 200 = 0$$

The term "dropped" corresponds to the root $-10/200$.

8 Unstable Equations

If an equation is not stable it can be made so by diminishing all the roots by a conveniently chosen number. Thus consider the equation

$$x^3 - 4x^2 + 6x - 4 = 0 \quad \ldots\ldots\ldots\ldots\ldots [19]$$

Since there are negative coefficients in Equation [19] this equation is not stable. We diminish the roots by 5, using conventional synthetic division as shown

				|5
1	−4	6	−4	
	5	5	55	
1	1	11	51	
	5	30		
1	6	41		
	5			
1	11			

Normally, the entry 1 is not repeated except at the end of the division. We have first divided synthetically left-to-right by 5, then the remainders by 5 again, and so on, until a division yields two remainders only. We form the equation

$$x^3 + 11x^2 + 41x + 51 = 0 \quad \ldots\ldots\ldots\ldots\ldots [20]$$

from remainders of the division. The roots of Equation [20] are equal to the roots of Equation [19] diminished by 5. Since $11 \times 41 > 51$ Equation [20] is stable. Equation [20] can be solved by the author's method. The roots are $-3, -4 \pm j$. It follows that the roots of Equation [19] are $2, 1 \pm j$.

9 Poor Convergence

With the methods described one may run into poor convergence or no convergence if the equation is on the border of stability. Consider the equation

$$x^4 + 2.2x^3 + 3.4x^2 + 2.4x + 2 = 0 \quad \ldots\ldots\ldots\ldots [21]$$

The roots of Equation [21] are $-0.1 \pm j, -1 \pm j$. The Routh array is now

1	3.4	2
2.2	2.4	
(3.4 − 1.1)	2	
(2.4 − 1.9)		

The numbers in the first column are written as differences in the way that they occur in the computation of the array. To avoid highly oscillatory dominant roots it is normally desirable to have the differences in the first column of the Routh array of the form $(a - b)$ where $a \geq 2b$. The difference $(2.4 - 1.9)$ in the array does not satisfy this requirement. Diminishing the roots of Equation [21] by 1 we obtain the division

					|1
1	2.2	3.4	2.4	2	
	1	3.2	6.6	9	
	3.2	6.6	9	11	
	1	4.2	10.8		
	4.2	10.8	19.8		
	1	5.2			
	5.2	16			
	1				
1	6.2				

The roots of the equation

$$x^4 + 6.2x^3 + 16x^2 + 19.8x + 11 = 0 \quad \ldots\ldots\ldots [22]$$

are thus one less than the corresponding roots of Equation [21], and are $-1.1 \pm j, -2 \pm j$. The Routh array is now

1	16	11
6.2	19.8	
(16 − 3.2)	11	
(19.8 − 5.3)		

The differences now satisfy the requirement $a \geq 2b$. In control design the roots $-1.1 \pm j$ are much to be preferred to $-0.1 \pm j$, which corresponds to a very oscillatory solution where a transient has many cycles before it dies out. In fact, for dominant roots (roots with numerically smallest real parts) $-a \pm bj$ it is desirable to have

$$2a \geq b$$

If Equation [14] is stable with all complex roots, and the author's right-to-left division starting with

$$x^2 + \frac{a_1}{2}x + \frac{a_2}{3}$$

does not converge rapidly to the solution, *diminishing the roots* of Equation [14] by a_1 *will always give an equation for which the convergence is rapid.*

10 Criticalness

The reader should be warned that some equations are critical. Consider thus the equation

$$x^4 + 4x^3 + 8x^2 + 8x + 4 = 0 \quad [23]$$

The left member factors into

$$(x^2 + 2x + 2)^2$$

Now make a slight change in the coefficients of Equation [23] and consider

$$x^4 + 4\frac{1}{3}x^3 + 8\frac{1}{3}x^2 + 8x + 4 = 0 \quad [24]$$

The left side of Equation [24] factors into

$$(x^2 + 3x + 3)\left(x^2 + \frac{4}{3}x + \frac{4}{3}\right)$$

Thus a slight change in the coefficients of an equation may result in a big change in the factors. This may mean a big change in transient performance, a factor to be considered in design.

11 Design of an Aircraft Gas-Turbine Governor

The problem of this section arose in the author's practice. From the design information supplied by the manufacturer and frequency-response runs, the author obtained the differential equation

$$\frac{d^2n}{dt^2} + 10\frac{dn}{dt} + n = 24{,}000z(t - 0.1) \quad [25]$$

relating fuel-valve position z in inches and gas-turbine (aircraft) rpm n, measured as deviations from equilibrium; i.e., a steady state where the valve and rpm are constant. For equilibrium (in this example zero mph, sea level at 7700 rpm) $n = z = 0$. Time t is measured in seconds. The expression $z(t - 0.1)$ represents a dead time of 0.1 sec, due to a combustion lag. Equation [25] can be written as

$$n = \frac{24{,}000e^{-0.1D}}{(0.1D + 1)(10D + 1)}z \quad [26]$$

for $D = d/dt$. The factor $(0.1D + 1)$ corresponds to a fuel-line time constant of 0.1 sec, the factor $(10D + 1)$ with a 10-sec time constant to engine damping, $e^{-0.1D}$ to the dead time of 0.1 sec.

The governor design selected was of the Woodward PG-type, whose equation is of the form

$$z = \frac{-K(D + A)}{D(D + B)}n \quad [27]$$

We shall take a governor where K is in the range

$$0.003{,}5 \leqq K \leqq 0.07$$

The engine rpm (deviation) n is the input to the governor which controls the fuel-valve position z. The fuel-valve position z in turn controls the engine rpm n.

We first neglect the small lags in Equation [26] and the engine damping, and take

$$n = \frac{2400}{D}z \quad [28]$$

Combining Equations [27] and [28] we have

$$(D^3 + BD^2 + 2400KD + 2400AK)n = 0 \quad [29]$$

The corresponding algebraic equation is

$$x^3 + Bx^2 + 2400Kx + 2400AK = 0 \quad [30]$$

By Routh's test for stability we must have

$$2400KB > 2400AK$$

whence

$$B > A$$

Because of practical considerations we take the roots in the form

$$-\alpha, \quad -\alpha \pm \alpha j \quad [31]$$

where α is as large as possible. In this case the cubic is

$$x^3 + 3\alpha x^2 + 4\alpha^2 x + 2\alpha^3 = 0 \quad [32]$$

It follows that we wish to choose K so that $2400K$ is as large as possible. This is the case if $K = 0.07$ whence

$$4\alpha^2 = 2400 \times 0.07 = 168 \quad [33]$$

Then

$$\alpha = 6.5$$

It follows that

$$A = 3.3, \quad B = 20 \quad [34]$$

Equation [30] now has the roots

$$-6.5, \ -6.5 \pm 6.5j \quad [35]$$

These roots correspond to faster transients than, we know from experience, we can expect to realize physically. Further, the roots are so "fast" that the neglected lags come into the picture to modify it seriously. We, therefore, lower the gain K. We arbitrarily try $2400K = 21$ since this decreases K substantially. Then

$$K = 0.008{,}75$$

whence Equation [30] becomes

$$x^3 + Bx^2 + 21x + 21A = 0 \quad [36]$$

Identifying Equation [36] with Equation [32] we have

$$\alpha = 2.3$$

and the roots of Equation [36] are now

$$-2.3, \ -2.3 \pm 2.3j \quad [37]$$

The governor equation is now

$$z = \frac{-0.008{,}75(D + 1.15)}{D(D + 6.9)}n \quad [38]$$

We shall check the effect of neglected quantities on the roots. We replace $e^{-0.1D}$ by

$$(-0.1D + 1)$$

whence in place of Equation [26] we have

$$n = \frac{24{,}000\,(-0.1D + 1)}{D^2 + 10D + 1}z \quad [39]$$

Eliminating z from Equations [38] and [39] we have

$$\{D(D^2 + 10D + 1)(D + 6.9) + 210(D + 1.15)(-0.1D + 1)\}n = 0$$

The corresponding algebraic equation is

$$x^4 + 17x^3 + 49x^2 + 190x + 240 = 0 \quad \ldots\ldots [40]$$

The solution is

1	17	49	190	240	|14
		37	173		
1	2.6	12	17		|1.6
	1.6	1.6	16		
1	1	10			

where the roots -14 and -1.6 are removed by right-to-left and left-to-right division respectively, and

$$x^2 + x + 10 = 0$$

is the reduced equation. The roots of Equation [40] are thus

$$-14, \quad -1.6, \quad -0.5 \pm 3j$$

In taking neglected factors into account the roots [37] have thus gone into

$$-1.6, \quad -0.5 \pm 3j \quad \ldots\ldots [41]$$

The complex roots in the set [41] correspond to highly oscillatory transients. It will therefore be necessary to modify the governor constants somewhat to improve the roots.

If we combine the governor Equation [27] with the turbine Equation [39] we obtain ($K = 0.008,75$)

$$\{D^4 + (10 + B)x^3 + (10B - 20)D^2 + (210 - 21A + B)D + 210A\}n = 0 \quad \ldots\ldots [42]$$

If B is large, say 100, the coefficients of Equation [42] become

1	110	980	$(310 - 21A)$	$210A$

Dropping the first two coefficients (which correspond to roots near -100 and -10) we obtain the quadratic equation

$$x^2 + (0.32 - 0.021A)x + 0.21A = 0$$

For any A ($0.32 > 0.021A > 0$) the coefficients of x and the constant are small and the roots correspond to very slow and unacceptable transients. It follows that B should not be large. For stability the coefficients in Equation [42] must all be positive. The coefficient of D^2 yields

$$B > 2$$

Thus B cannot be too small.

The coefficient of x^3 in Equation [42] is large for each B. We, therefore, drop the first term in this equation to obtain

$$\left\{D^3 + \frac{10B - 20}{10 + B} D^2 + \frac{210 - 21A + B}{10 + B} D + \frac{210A}{10 + B}\right\} n = 0$$

With $B = 10$ this yields the coefficients

1	4	$(11 - A)$	$10A$

and with $B = 20$ we have

1	6	$7.7 - 0.7A$	$7A$

For $B = 10$, $A = 1$ we have (left-to-right division)

1	4	10	10	|1.6
	1.6	3.8	10	
1	2.4	6.2		

For $B = 20$, $A = 1$ with right-to-left division we have

1	6	7	7	4.9
		5.6		
1	1.1	1.4		

For $B = 10$, $A = 0.5$ we have by left-to-right division

1	4	10	5	|0.6
	0.6	2	4.8	
1	3.4	8		

Finally with $B = 20$, $A = 0.5$ and right-to-left division we have

1	6	7.4	3.5	|4.6
		6.6		
	1.4	0.76		

We thus have the roots

A	B	Real root	Imaginary roots
0.5	10	-0.6	$-1.7 \pm 2.2j$
1		-1.6	$-1.2 \pm 2.2j$
0.5	20	-4.6	$-0.7 \pm 0.5j$
1		-4.9	$-0.55 \pm j$

The complex roots for $A = 1$ in the foregoing table have relatively large imaginary parts, compared to the real parts, whereas the dominant roots (roots with numerically small real parts) for $A = 0.5$ are small numerically. We wish to have real and imaginary parts of the complex roots approximately equal, so that these roots will not correspond to transients that are too oscillatory, and we do not wish these roots to be small in magnitude, so as to make the transients slow. Our objectives can be achieved by a choice of A between 0.5 and 1. The complex roots for $B = 10$ have relatively large imaginary parts, whereas for $B = 20$ the complex roots have small real parts, and correspond to slow transients. We, therefore, choose a value of B between 10 and 20. We take

$$A = 0.8$$
$$B = 15$$
$$K = 0.008,75$$

The corresponding algebraic equation is

$$x^4 + 25x^3 + 130x^2 + 210x + 170 = 0 \quad \ldots\ldots [43]$$

Since the coefficient 25 of x^3 in this equation is large compared to the coefficient of x^4, and in turn the coefficient 130 of x^2 is large compared to the coefficient 25 of x^3, we may drop the x^4 and x^3 terms of this equation. The resulting quadratic has roots equal to

$$-0.8 \pm 0.8j$$

approximately. These roots are satisfactory. We, therefore, solve Equation [43] more precisely. The solution follows (right-to-left division) where we round off numbers to two digits.

1	25	130	210	170	|19
		119	201		
1	6.3	11	9		|4.2
		8.9			
1	2.1	2.1			

We have removed successfully the roots -19, -4.2 and obtained the reduced quadratic

$$x^2 + 2.1x + 2.1 = 0$$

The roots of Equation [43] are thus (approximately)

$$-19, \quad -4.2, \quad -1 \pm j$$

These roots are satisfactory.

To be sure that the approximation employed for $e^{-0.1D}$ was valid we try the more accurate approximation

$$0.005D^2 - 0.1D + 1$$

obtained from the expansion

$$e^{-0.1D} = 1 - 0.1D + \frac{(0.1D)^2}{2!} - \frac{(0.1D)^3}{3!} + \ldots\ldots \quad [44]$$

The turbine equation is now

$$n = \frac{24{,}000\,(0.005D^2 - 0.1D + 1)}{D^2 + 10D + 1}\,z \ldots\ldots\ldots [45]$$

Combining Equations [45] and [27] with

$$K = 0.008{,}75, \quad A = 0.8, \quad B = 15$$

we obtain

$$(D^4 + 26D^3 + 130D^2 + 210D + 170)n = 0$$

This equation is practically identical with Equation [43] and the dominant roots

$$-1 \pm j$$

are not changed. Replacing $A = 1.15$ by 0.8 and $B = 6.9$ by 15 we have transformed the roots

$$-14, \; -1.6, \; -0.5 + 3j, \; -0.5 - 3j$$

into

$$-19, \quad -4.2, \quad -1 + j, \quad -1 - j$$

respectively.

The author's method of finding complex roots also can be used for solving for real roots by using the first three terms in Equation [5] to obtain a trial divisor, even though the corresponding roots are real. Thus to solve Equation [43] we can use the trial divisor

$$x^2 + 20x + 100$$

suggested by the first three terms of this equation. Right-to-left divisions then rapidly lead to the divisor

$$x^2 + 23x + 80$$

of the left side of Equation [43].

It is the experience of the author that by neglecting the proper factors one always can reduce the differential equation relating the input and output of a physical system to an equation of low order, such as the second or third.

The example of this section did not involve an equation with four or more complex roots. The design approach given here extends directly to such examples.

BIBLIOGRAPHY

1 "Frequency Response Symposium," Trans. ASME, vol. 76, 1954, pp. 1145–1393. This issue was entirely devoted to frequency response.

2 "Control System Dynamics," by W. R. Evans, McGraw-Hill Book Company, Inc., New York, N. Y., 1954.

3 "Theory of Equations," by J. V. Uspensky, McGraw-Hill Book Company, Inc., New York, N. Y., 1948, pp. 318–331.

4 "Dynamics of a System of Rigid Bodies," by E. J. Routh, Macmillan and Company, Ltd., London, England, 1877.

5 "The Use of Zeros and Poles for Frequency Response or Transient Response," by W. R. Evans, discussion by Rufus Oldenburger, Trans. ASME, vol. 76, 1954, pp. 1340–1343.

6 "Mathematical Engineering Analysis," by Rufus Oldenburger, The Macmillan Company, New York, N. Y., 1950.

7 "Practical Computational Methods in the Solution of Equations," by Rufus Oldenburger, *American Mathematical Monthly*, vol. 45, June-July, 1948, pp. 334–342.

8 "Method of Successive Approximations of Evaluating the Real and Complex Roots of Cubic and Higher Order Equations," by Shih-Nge Lin, *Journal of Mathematics and Physics*, vol. 20, August, 1941, pp. 231–242.

9 "Mathematical Methods in Engineering," by T. von Karman and M. A. Biot, McGraw-Hill Book Company, Inc., New York, N. Y., 1940, p. 247.

Appendix

We consider the stable quartic

$$x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0 \ldots\ldots\ldots [14]$$

with all complex roots. We assume that the left side factors into

$$(x^2 + A_1x + A_2)(x^2 + B_1x + B_2) \ldots\ldots\ldots [46]$$

for numbers A_1, A_2, B_1, and B_2. We may suppose that the numbers are so labeled that

$$A_1 \geq B_1 \ldots\ldots\ldots [47]$$

Multiplying out the Polynomial [46] we obtain the form

$$x^4 + (A_1 + B_1)x^3 + (A_2 + A_1B_1 + B_2)x^2 + (A_2B_1 + A_1B_2)x + A_2B_2 = 0 \ldots\ldots [48]$$

of Equation [14]. We shall say that a number A dominates a number B, written $A \gg B$ or as $B \ll A$ when $A > 5B$.

Suppose that A_2 is not dominated by $(A_1B_1 + B_2)$, written

$$A_2 <\!/\!< A_1B_1 + B_2$$

and meaning

$$5A_2 \geq A_1B_1 + B_2$$

The quadratic

$$x^2 + (A_1 + B_1)x + (A_2 + A_1B_1 + B_2)$$

formed from the first three terms in Equation [48] is then, at least roughly, approximately the same as the factor

$$x^2 + A_1x + A_2 \ldots\ldots\ldots [49]$$

of the quartic on the left of Equation [14]. Here the author takes $\alpha/6$ to be "approximately" equal to α in the sense that 100 is a much better approximation to 600 than 1 is to 600. A rough approximation is better than none at all.

The first row of Relations [15] corresponds to this case where α and β are to be identified with A_1 and A_2, respectively, as will now be proved. We note that Relation [47] yields

$$2A_1 \geq A_1 + B_1$$

whence

$$2\alpha \geq a_1$$

Thus

$$\frac{a_1}{2} \leq \alpha < a_1$$

From

$$5A_2 \geq A_1B_1 + B_2$$

we have

$$6A_2 \geq A_2 + A_1B_1 + B_2$$

whence

$$6\beta \geq a_2$$

It follows that

$$\frac{a_2}{6} \leqq \beta < a_2$$

We have thus derived the first row of Relations [15].

Suppose now that

$$A_1B_1 + B_2 \gg A_2 \quad \ldots\ldots [50]$$

Since the roots of Equation [14] are all complex we have

$$A_1^2 < 4A_2, \quad B_1^2 < 4B_2 \quad \ldots\ldots [51]$$

If $B_2 \leqq A_2$ and Inequality [50] holds, then

$$A_2 + A_1B_1 \gg A_2 \quad \ldots\ldots [52]$$

From the definition of dominance we have

$$A_1B_1 > 4A_2 \quad \ldots\ldots [53]$$

Since the Relation [47] holds we have

$$A_1^2 > 4A_2$$

contradicting Inequalities [51]! Thus the Dominance [52] does not hold in any case, and we have

$$B_2 > A_2 \quad \ldots\ldots [54]$$

Suppose that

$$B_2 \ll A_1B_1 + A_2 \quad \ldots\ldots [55]$$

holds as well as the Dominance [50]. In view of Inequality [54] and Relation [55] we have the Dominance [52]. Since this dominance does not hold we have

$$B_2 <\!/\!< A_1B_1 + A_2 \quad \ldots\ldots [56]$$

The coefficient of x^2 in Equation [48] can now be approximated (at least in a rough sense) by B_2.

If $A_1B_2 \ll A_2B_1$ the Relation [47] implies that

$$B_2 \ll A_2$$

This contradicts Relation [56]. Thus we have

$$A_1B_2 <\!/\!< A_2B_1 \quad \ldots\ldots [57]$$

and the coefficient of x in Equation [48] can be replaced by A_1B_2. The last three terms of the left side of Equation [48] can thus be approximated by

$$B_2x^2 + A_1B_2x + A_2B_2$$

yielding the Factor [49] of the left side of Equation [48].

Relation [56] implies that

$$5B_2 \geqq A_1B_1 + A_2$$

or

$$6B_2 \geqq A_2 + A_1B_1 + B_2$$

which means that

$$B_2 \geqq \frac{a_2}{6}$$

From Relation [48] we have $a_2 > B_2$, whence

$$a_2 > B_2 \geqq \frac{a_2}{6} \quad \ldots\ldots [58]$$

Relation [57] implies that

$$5A_1B_2 \geqq A_2B_1$$

or

$$6A_1B_2 \geqq A_2B_1 + A_1B_2 = a_2$$

whence in view of Relation [48]

$$a_2 > A_1B_2 \geqq \frac{a_2}{6} \quad \ldots\ldots [59]$$

From the constant terms in Equations [14] and [48] we have

$$B_2 = a_4/A_2 \quad \ldots\ldots [60]$$

Relation [58] now yields (remembering that β is the same as A_2)

$$\frac{a_4}{a_2} < \beta \leqq \frac{6a_4}{a_2} \quad \ldots\ldots [61]$$

given in the third line of Relations [15]. From Relations [59] and [60] we have

$$\frac{a_2}{6a_4} \leqq \frac{\alpha}{\beta} < \frac{a_2}{a_4}$$

as in the third row of Relations [15].

The method of proof used for quartics generalizes to higher-degree equations. If the roots of Equation [48] are not all almost real the inequality $A > 5B$ used to define dominance can be changed to $A > 4B$.

Discussion

HAROLD CHESTNUT.[4] A worthwhile contribution to the control system literature has been made by Professor Oldenburger with the presentation of this paper. Although the use of Routh's Criterion and the left-to-right synthetic division methods of determining system stability had been used for some time past, the right-to-left synthetic division method, the quadratic method of factoring out complex roots, and the general algebraic approach to design of automatic controls have not received much attention as being desirable design techniques. Professor Oldenburger has done a commendable job of demonstrating ways of estimating the values of the largest and the smallest roots of the system, and then of determining the values of these and the remaining roots for fairly complex control-system equations. The effectiveness of algebraic methods has for some time now been minimized, yet this method does provide a quick analytical tool for many control problems.

In keeping with the conference theme, "Application of New Control Analysis Techniques," I should like to suggest an alternate design approach which is directed at determining the roots of a control system using the open-loop attenuation characteristic as a starting point. This method is one that was described by Kan Chen at the 1957 AIEE Winter Meeting and is contained in AIEE Paper No. 57-182, "A Quick Method for Estimating Closed-Loop Poles of Control Systems." It is an extension of work of George Biernson and has the advantages of being able to handle systems of considerable complexity without an ever-increasing degree of effort. It is hoped that Chen will also provide a discussion to this paper that will present the salient features of his method.

As a direct question of Professor Oldenburger, I should appreciate learning from him his comments on the relative merits of the algebraic versus other methods such as frequency response or root locus means for system analysis. In what type problems would one method be preferable to another?

[4] Engineer, General Electric Company, Schenectady, N. Y.

KAN CHEN.[5] I wish to join Mr. Chestnut in commending Professor Oldenburger for having made a remarkable contribution to the control-system literature with the presentation of this paper. Since Mr. Chestnut referred to a method suggested by myself as an alternative design approach to Professor Oldenburger's method, I shall describe my method briefly, illustrate its use by working out a problem in this paper, and express my opinions on the relative merits of the two methods.

The method I proposed is an extension of Mr. Biernson's work (10, 11)[6] and is based on a comprehensive use of the root locus plot (12) and the frequency asymptote plot (13). To fix ideas, let us denote the open-loop transfer function of a unity-feedback system by $G(D)$, the input of the system by R, the error by E, and the output by C. Then, for $|G| \leqslant -15$ db, the following approximation may be written

$$\frac{C(D)}{R(D)} = \frac{G(D)}{1 + G(D)} \approx G(D) \quad \ldots\ldots [62]$$

Under this condition, it is clear that the closed-loop poles of $C(D)/R(D)$ are approximately the open-loop poles of $G(D)$. For $|G| \gtrsim +15$ db, the following approximation may be written

$$\frac{E(D)}{R(D)} = \frac{1}{1 + G(D)} \approx \frac{1}{G(D)} \quad \ldots\ldots [63]$$

Under this condition, it is clear that the closed-loop poles of $C(D)/R(D)$, which are the same as the poles of $E(D)/R(D)$, are approximately the open-loop zeros of $G(D)$.

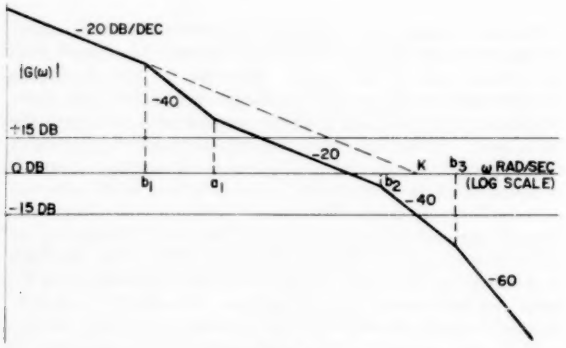


FIG. 1

Thus if the open-loop frequency asymptote, $|G(\omega)|$ in db vs ω in logarithmic scale, is like that shown in Fig. 1, then two approximate closed-loop poles are at $-a_1$ and $-b_3$. This is deduced by inspection that the break at a_1 corresponds to an open-loop zero and occurs above the +15-db line and that the break at b_3 corresponds to an open-loop pole and occurs below the −15-db line. Of course, it is a very rough approximation to consider the two closed-loop poles to be right at $-a_1$ and $-b_3$. However, since the closed-loop pole near $-b_3$ corresponds to a rapidly decaying transient term, it will not affect the over-all system transient significantly to consider one of the closed-loop poles to be right at $-b_3$. The closed-loop pole near $-a_1$ corresponds to a slowly decaying transient term but is not a dominant pole because of its proximity to the closed-loop zero at $-a_1$. The approximate distance between the closed-loop pole and zero near $-a_1$ can be shown (14) to be inversely proportional to the frequency-asymptote gain at $\omega = a_1$.

The dominant closed-loop poles can be approximately determined from the gain and break points of G within the ±15-db band only. In other words, the breaks outside the ±15-db band may be ignored in this case. For the system with the open-loop frequency asymptotes given in Fig. 1, the transfer function $G(D)$ is now approximated by

$$G(D) \approx \frac{K/b_3}{D(D + b_2)} \quad \ldots\ldots [64]$$

and

$$\frac{G(D)}{1 + G(D)} \approx \frac{K/b_3}{K/b_3 + D(D + b_2)} \quad \ldots\ldots [65]$$

The original quartic characteristic equation has thus been reduced to a quadratic equation which can be easily solved. When a closed-loop system has satisfactory stability, the simplification of $G(D)$ in the manner described above will usually result in a transfer function not higher than the third order (14). The approximate dominant closed-loop poles can, therefore, be solved either algebraically or graphically with the aid of a few generalized root-loci charts (14).

Now let us use this method to solve the problem in Professor Oldenburger's paper for the following set of values

$A = 0.8$
$B = 15$
$K = 0.00875$

The corresponding open-loop transfer function is

$$G(D) = \frac{210(D + 0.8)e^{-0.1D}}{D(D + 15)(0.1D + 1)(10D + 1)} \quad \ldots\ldots [66]$$

In order to use the suggested method, the transfer function $G(D)$ has to be stable and minimum-phase. Thus we use the following approximation

$$e^{-0.1D} \approx \frac{1}{1 + 0.1D} \quad \ldots\ldots [67]$$

This will make

$$G(D) \approx \frac{11.2\left(\frac{D}{0.8} + 1\right)}{D\left(\frac{D}{0.1} + 1\right)\left(\frac{D}{15} + 1\right)\left(\frac{D}{10} + 1\right)^2} \quad \ldots [68]$$

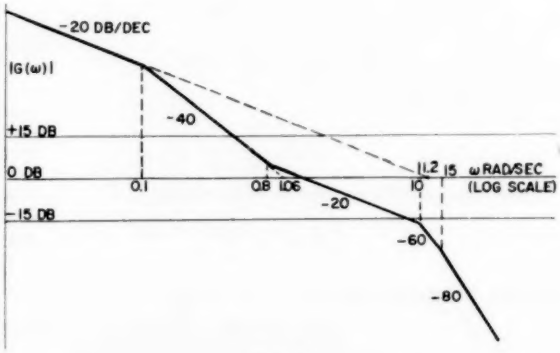


FIG. 2

[5] New Products Engineering Department, Westinghouse Electric Corporation, Pittsburgh, Pa.

[6] Numbers in parentheses refer to the Bibliography at the end of the discussion.

which is fifth order. The corresponding frequency asymptote plot is shown in Fig. 2. The double break at 10 and the single break at 15 are below the −15-db line. Thus three approximate closed-loop poles are

$$-10,\ -10,\ \text{and}\ -15$$

Within the ±15-db band, the transfer function

$$G(D) \approx \frac{(1.06)^2\left(\frac{D}{0.8}+1\right)}{D^2} \quad \ldots\ldots\ldots\ldots [69]$$

and

$$\frac{G(D)}{1+G(D)} \approx \frac{(1.06)^2\left(\frac{D}{0.8}+1\right)}{D^2+(1.06)^2\left(\frac{D}{0.8}+1\right)} \quad \ldots\ldots\ldots [70]$$

Solving the quadratic characteristic Equation [70] yields the following dominant closed-loop poles

$$-0.7 \pm j\,0.8$$

Comparing these results with those obtained by Professor Oldenburger, we see that the solution yielded by this method is not as accurate, but the salient features of transient indicated by the solution are the same as those indicated by Professor Oldenburger's solution. That is, the damping ratio is approximately unity; the delay time contributed by the poles at −19 and −4.2 given by the author's solution is about the same as that contributed by the poles at −10, −10, and −15 given by the discusser's solution (15).

In the light of these results, I shall venture my opinion on the relative merits of the two methods as follows:

(1) Professor Oldenburger's method yields better accuracy by providing a reiteration procedure for improving accuracy. The reiteration procedure may be shortened by exercising engineering judgment. On the contrary, in the method I proposed, improved accuracy of solution is sacrificed for the sake of simplicity, resulting in complete elimination of reiteration processes. It is felt that the accuracy given is good enough for design purposes and that the job of doing accurate calculation should be left for a computer.

(2) Using the Bode plot as a working medium, the method I proposed seems to provide a better insight to the design problem. It has been demonstrated (14) that the straightforward design procedure advocated by Professor Truxal (16) can be conveniently carried out with the use of my method. However, this advantage of facilitating system design is limited, at least at present, to single-loop control systems only. On the contrary, Professor Oldenburger's method is obviously as valid for multi-loop as for single-loop control systems.

Being more familiar with my own method than with Professor Oldenburger's method, I cannot help being somewhat subjective in the comparison. I trust Professor Oldenburger will also comment on the relative merits of the various methods in his closure of this paper.

BIBLIOGRAPHY

10 "Quick Methods for Evaluating the Closed-Loop Poles of Feedback Control Systems," by G. A. Biernson, Trans. AIEE, vol. 72, part II, 1953, pp. 53–70.

11 "A General Technique for Approximating Transient Response From Frequency Response Asymptotes," by G. A. Biernson, Trans. AIEE, vol. 75, part II, 1956, pp. 253–273.

12 "Control-System Dynamics," by W. R. Evans, McGraw-Hill Book Co., Inc., New York, N. Y., 1954.

13 "Servomechanisms and Regulating Systems," by H. Chestnut and R. W. Mayer, vol. 1, John Wiley & Sons, Inc., New York, N. Y., 1951, pp. 415–416.

14 "A Quick Method for Estimating Closed-Loop Poles of Control Systems," by K. Chen, Trans. AIEE, Paper 57–182.

15 "Quasi-Linearization Techniques for Transient Study of Nonlinear Feedback Control Systems," by K. Chen, Appendix, Trans. AIEE, vol. 74, part II, 1955, pp. 361–363.

16 "Automatic Feedback Control System Synthesis," by J. G. Truxal, Chs. 5 and 6, McGraw-Hill Book Co., Inc., New York, N. Y., 1955.

Author's Closure

The characteristic equations for control systems are generally such that in a first approximation one can drop the higher degree terms and reduce these equations to the fifth or lower degree, which can be readily manipulated. This justifies the simplification of a frequency-response locus as illustrated by Mr. Chen, so that for minumum phase systems one can draw satisfactory conjectures as to the nature of closed-loop response from that for open loop. Knowledge of the relationship between open-loop poles and zeros and closed-loop poles, brought out by Mr. Chen, is an invaluable aid to understanding control systems.

The frequency-response approach is indicated for various problems that cannot be studied rapidly by algebraic techniques, as is sometimes the case for systems with distributed constants. The frequency-response approach often enables one to determine quickly the cause of hunting of a physical system in operation. It is invaluable for obtaining and checking the transfer functions of system components.

The root-locus method is of considerable value in giving a geometric picture of the relation between closed loop poles and properties of the open loop such as gain.

As normally applied the frequency-response and root-locus approaches involve the graphing of curves. The writer has worked many problems by the methods of Evans, Chen, Bode, and others. Although he has used these other methods often, by algebraic techniques he generally in much less time has been able to solve the day-by-day design problems that arise in industry. There are always severe limitations on the design or redesign of control equipment. Often one cannot change the gain of the open loop, arbitrarily insert leads and lags, and carry out other standard design techniques which are fairly easy to perform in the area of electrical communication. The parameters generally enter in a complicated way, and this is true of the other properties that may be modified. *No approach can give a deeper or more thorough insight than the analytical with system parameters kept as such. The writer believes that analytical (as opposed to graphical) methods should be employed where possible.*

Although in a rough study accuracy is not needed, it is desirable to attain it on paper where it can be done quickly, as by the writer's solution method. Running to a computer is time-consuming even if it is in the same room. Although in industry the writer had an excellent electronic computer immediately available with a staff ready to give assistance, he found paper solutions more efficient in first approximation studies to linear problems. A computer or graph will handle only one numerical problem at a time, or class of equivalent numerical cases.

A control expert must familiarize himself with as many solution methods as possible. Only by using each method on several problems can he develop adequate computational facility in the method. With a large kit of theoretical tools at his disposal he can select the method of solution that fits the problem.

The writer's method of solving equations applies to all algebraic equations with real coefficients, not only stable equations, and can therefore be used wherever such algebraic equations are employed in science and industry.

Statistical Treatment of Sampled-Data Control Systems for Actual Random Inputs

By MASAHIRO MORI,[1] CHIBA CITY, JAPAN

Fundamental statistical relations of sampled-data systems in terms of the correlation functions of time series, the pulse spectral densities, and the modified z-transform are presented. If one is interested in signals between sampling instants, the modified z-transform method can be applied easily to the statistical treatments, as are presented in the paper. The relation between this method and the analysis on the basis of variable-system theory is proved. Examples are given discussing discrete compensation of sampled-data control system from the statistical standpoint.

Nomenclature

The following nomenclature is used in the paper:

- c = output or controlled variable
- f = frequency
- f_0 = sampling frequency
- G = continuous transfer function of system whose impulsive response is $g(t)$
- G^* = pulse transfer function or z-transform of $g(t)$
- g = wave form
- h = integer identifying sampling instants
- K = system function
- k = integer identifying sampling instants
- $\mathfrak{L}$ = symbol for Laplace transformation
- M = mean square of absolute value of modified pulse-frequency transfer function with respect to m
- m = $1 - \Delta$
- N = integer
- n = integer identifying sampling instants
- q = integer identifying sampling instants
- R^* = correlation function of time series
- Re = symbol for real part
- r = input or reference input
- S = continuous spectral density
- S^* = pulse spectral density
- s = $\sigma + j\omega$, operator of Laplace transformation
- T = $1/f_0$, sampling period
- t = time
- u = disturbance
- W^* = pulse transfer function from reference input to error
- $\mathcal{Z}$ = symbol for z-transformation
- z = e^{sT}, operator of z-transformation
- Δ = dead time measured as a fraction of a sampling period
- δ_T = train of unit impulses
- ϵ = error
- ω = $2\pi f$, angular frequency
- ω_0 = $2\pi f_0$

[1] Institute of Industrial Science, University of Tokyo.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-10.

Introduction

It is noted that sampled-data control systems which attract attention in the field of servomechanisms (1, 2)[2] have considerable application also in industrial process control as follows: Sampled-data systems make it possible to incorporate digital computers (7) in process control, since their input and output are of intermittent form by nature. And in sampled-data systems it is possible to take advantage of multiplex control methods to control a number of processes by one controller with scanners or with multipoint switching mechanisms. Moreover, by means of discrete compensating units using sampled data (4, 5) it is possible to compensate a process to have an over-all desirable transfer function beyond the ability of continuous compensation.

Though sampled-data control systems have been studied by many researchers by the indicial-response method, the author considers that this method is not appropriate to treat sampled-data systems for the following reasons: (*a*) The sampled-data control system cannot work satisfactorily without the assumption that the highest frequency component of its input signal must be lower than one half of the sampling frequency (6, 7), while frequency components of the step (or the ramp) input expand to infinity beyond this limit. This is a contradiction. (*b*) Characteristic of the hold circuit which is always involved in the system and in essence smoother extrapolator of sampled data is highly dependent upon the mutual relation between order of the hold circuit and its input wave forms. For instance, a zero-order hold circuit has an excellent characteristic of extrapolation only for step-type-input wave form and poor characteristics for other input wave forms. Almost all actual input wave forms are random. Accordingly, analysis and synthesis in terms of the system response to special test signals such as indicial-response method are not appropriate for the sampled-data systems. Therefore the author proposes a treatment of sampled-data control systems for actual random inputs.

The first part of the paper is a brief review of mathematical relations. Then, fundamental formulas for statistical evaluations of mean-square values both of sampled and of total-output signals are presented. Use of modified z-transform for statistical calculation of sampled-data systems based upon variable system theory constitutes the final part of the paper.

Background of z-Transform

Basic relations of sampled-data systems will be reviewed briefly. For the analysis and synthesis of sampled-data systems, as is well known, the z-transform theory (1, 2) offers a useful method. The "z-transform" of a wave form $g(t)$ or the "pulse-transfer function" of a system whose impulsive response is $g(t)$ is given by the following two equivalent forms (8, 9)

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

$$G^*(z) = \sum_{n=0}^{\infty} g(nT)z^{-n} \quad [1]$$

$$G^*(z) = \frac{1}{T}\sum_{n=-\infty}^{\infty} G(s + jn\omega_0) \quad [2]$$

where $z = e^{sT}$; T, sampling period; s, operator of the Laplace transform; $G(s)$, Laplace transform of $g(t)$; $\omega_0 = 2\pi/T$ and n, integer.

With reference to Fig. 1, the pulse-transfer function $G^*(z)$ is the ratio of z-transforms of output signal $c(t)$ and input signal $r(t)$.

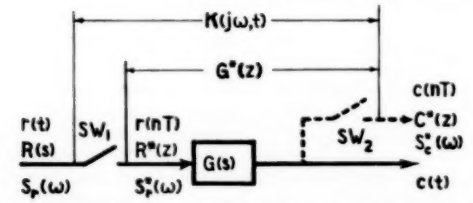


FIG. 1 PULSE-TRANSFER FUNCTION $G^*(z)$ AND SYSTEM FUNCTION $K(j\omega,t)$. SAMPLER SW_1 IS SYNCHRONIZED WITH SAMPLER SW_2

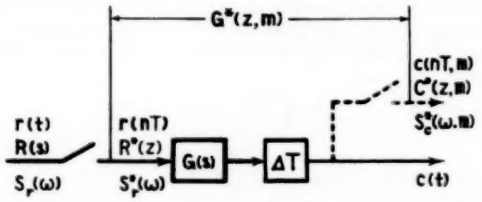


FIG. 2 FICTITIOUS DEAD TIME ΔT TO OBTAIN TOTAL RESPONSE BETWEEN SAMPLING INSTANTS BY MODIFIED z-TRANSFORM $m = 1 - \Delta$, $0 \leq \Delta < 1$

The analysis using the z-transform has the limitations of yielding the response only at sampling instants. However, to obtain the response between sampling instants, the z-transform can be modified by inserting a fictitious dead time ΔT in the forward path as is indicated in Fig. 2 (1, 3). One can obtain the total response between sampling instants by varying Δ from zero to unity. Pulse-transfer function of the system having the dead time ΔT is termed "modified pulse-transfer function." The modified pulse-transfer function or "modified z-transform" of a wave form $g(t)$ is given by the following two equivalent forms

$$G^*(z, m) = z^{-1}\sum_{n=0}^{\infty} g(nT + mT)z^{-n} = \sum_{n=0}^{\infty} g(nT - \Delta T)z^{-n} \quad [3]$$

$$G^*(z, m) = \frac{1}{T}\sum_{n=-\infty}^{\infty} G(s + jn\omega_0)e^{-(s+jn\omega_0)(1-m)T} \quad [4]$$

where $0 < m \leq 1$ and $m = 1 - \Delta$. Equation [4] derived by the author (derivation appears in proof of Equation [39] in Appendix 3) constitutes a basis for the statistical treatments of the present paper. A short list of z-transforms and modified z-transforms is given in Table 1.

By substituting $z = e^{j\omega T}$ in $G^*(z)$ and $G^*(z, m)$, for stable systems we can obtain "pulse-frequency transfer function" $G^*(e^{j\omega T})$ and "modified pulse-frequency transfer function" $G^*(e^{j\omega T}, m)$ as follows

$$G^*(e^{j\omega T}) = G^*(z)\big|_{z=e^{j\omega T}} \quad [5]$$

TABLE 1 SHORT LIST OF z-TRANSFORMS

Laplace transform	z-transform	Modified z-transform
$\frac{1}{s}$	$\frac{z}{z-1}$	$\frac{1}{z-1}$
$\frac{1}{s+a}$	$\frac{z}{z-e^{-aT}}$	$\frac{e^{-amT}}{z-e^{-aT}}$
$\frac{a}{s^2+a^2}$	$\frac{z\sin aT}{z^2-2z\cos aT+1}$	$\frac{z\sin amT+\sin(1-m)aT}{z^2-2z\cos aT+1}$

$$G^*(e^{j\omega T}, m) = G^*(z, m)\big|_{z=e^{j\omega T}} \quad [6]$$

where ω is angular frequency.

$G^*(z)$, $G^*(z, m)$, $G^*(e^{j\omega T})$, and $G^*(e^{j\omega T}, m)$ are periodic functions in frequency f, with period $1/T$.

CORRELATION FUNCTIONS AND SPECTRAL DENSITIES FOR TIME SERIES

Correlation functions and spectral densities for time series[2] are useful to treat sampled-data systems. Analogous relations hold for sampled-data systems as for continuous-data systems, which are presented in the following.

1 *Correlation Functions.* Autocorrelation function $R_r^*(k)$ of a time series $r(nT)$ and cross correlation function $R_{rc}^*(k)$ of time series $r(nT)$ and $c(nT)$ are defined as

$$R_r^*(k) = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N} r(nT)r(nT + kT) \quad [7]$$

$$R_{rc}^*(k) = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N} r(nT)c(nT + kT) \quad [8]$$

where k and n are integers. These are widely known.

2 *Pulse Spectral Densities.* We introduce an auxiliary time series $r_N(nT)$ as

$$\left.\begin{aligned} r_N(nT) &= r(nT) \quad \text{when } -N \leq n \leq N \\ &= 0 \quad \text{elsewhere} \end{aligned}\right\} \quad [9]$$

Then the definition of the spectral density $S_r^*(\omega)$ of the series $r(nT)$ is

$$A_N(e^{j\omega T}) = T\sum_{n=-N}^{N} r_N(nT)e^{-j\omega nT} \quad [10]$$

$$S_r^*(\omega) = \lim_{N\to\infty}\frac{1}{(2N+1)T}|A_N(e^{j\omega T})|^2 \quad [11]$$

The $S_r^*(\omega)$ is termed "pulse spectral density" in the paper.

Now, $A_N(e^{j\omega T})$ is a periodic function in frequency f with period of $1/T$ and is completely determined by values of ω in the range from $\omega = -\pi/T$ to $\omega = \pi/T$. It follows from the periodicity of $A_N(e^{j\omega T})$ and the orthogonality of the functions $e^{-j\omega T}$ that

$$T\sum_{n=-N}^{N}\{r_N(nT)\}^2 = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T}|A_N(e^{j\omega T})|^2 d\omega \quad [12]$$

Thus, the right side term of Equation [12] expresses the energy of the rectanglar waveform indicated in Fig. 3, by which the continuous waveform $r_N(t)$ is approximated. Therefore $|A_N(e^{j\omega T})|^2$ may be a measure of the energy in the frequency range ω to $\omega + d\omega$, of the series which has equal energy to the rectangular waveform. If we divide the energy spectrum by the effective time duration $(2N + 1)T$, we get the power spectrum. Finally, at

[2] Some of them and their mutual relation have been mentioned briefly in the Bibliography (2).

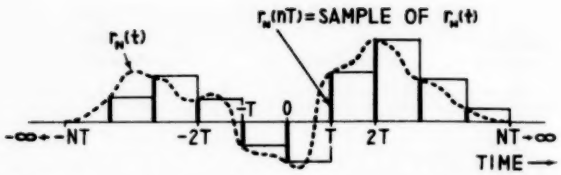


FIG. 3 ILLUSTRATION FOR PHYSICAL MEANING OF PULSE SPECTRAL DENSITY $S_r^*(\omega)$. CONTINUOUS WAVE FORM $r_N(t)$ IS APPROXIMATED BY A SET OF RECTANGLES

the limit of $N \to \infty$, the pulse spectral density is a measure of the frequency distribution of the power in the series having equal power to the rectangular wave form extended from $NT = -\infty$ to $NT = \infty$.

The following equation for mean-square value can be derived from Equation [12]

$$\overline{\{r(nT)\}^2} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \{r_N(nT)\}^2 = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} S_r^*(\omega) d\omega \quad \ldots\ldots [13]$$

It is possible to define "pulse cross-spectral density" for two time series $r(nT)$ and $c(nT)$ is an analogous way. Let

$$\left.\begin{aligned} A_N(e^{j\omega T}) &= T \sum_{n=-N}^{N} r_N(nT) e^{-j\omega nT} \\ \text{and} \qquad B_N(e^{j\omega T}) &= T \sum_{n=-N}^{N} c_N(nT) e^{-j\omega nT} \end{aligned}\right\} \quad \ldots\ldots [14]$$

The pulse cross-spectral density is then defined as

$$S_{rc}^*(\omega) = \lim_{N\to\infty} \frac{1}{(2N+1)T} \bar{A}_N(e^{j\omega T}) B_N(e^{j\omega T}) \quad \ldots [15]$$

where $\wedge$ is symbol for the conjugate.

3 *Mutual Relations.* As is known, the relation between correlation functions and spectral densities of continuous data is the Fourier transform, which is Fourier integral.

In sampled data, the corresponding relation between the autocorrelation function and the pulse spectral density mentioned previously is the Fourier series and its coefficients as follows

$$R_r^*(k) = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} S_r^*(\omega) e^{j\omega kT} d\omega \quad \ldots\ldots [16]$$

$$S_r^*(\omega) = T \sum_{k=-\infty}^{\infty} R_r^*(k) e^{-j\omega kT} \quad \ldots\ldots [17]$$

The same relation holds between the cross correlation and the pulse cross-spectral density.

The way to prove equations given in this section has been shown in Bibliography (2). The equations correspond to those for continuous data, as shown in Table 2 of Appendix 1.

Statistical Output-Input Relations

Analogous treatment for sampled-data systems as for continuous-data systems are presented in the following.

Let us start the following basic relation which can be proved. With reference to Fig. 2, if $G^*(e^{j\omega T}, m)$ is the modified pulse-frequency transfer function of a linear stable system whose continuous transfer function is $G(s)$, and $S_r^*(\omega)$ is the pulse spectral density of the input samples $r(nT)$ at sampling instants, then the pulse spectral density $S_c^*(\omega, m)$ of the output samples $c(nT, m)$ at the instants delayed ΔT from the sampling instants is given as

$$S_c^*(\omega, m) = |G^*(e^{j\omega T}, m)|^2 S_r^*(\omega) \quad \ldots\ldots [18]$$

The proof of Equation [18] appears in Appendix 2. The relation reduced to the simpler form

$$S_c^*(\omega) = |G^*(e^{j\omega T})|^2 S_r^*(\omega) \quad \ldots\ldots [19]$$

for (a special case) pulse spectral density $S_c^*(\omega)$ of the output samples $c(nT)$ at the sampling instants, Fig. 1.

Now, by applying the relation of Equation [13] to the output, the mean-square value of the output samples $c(nT, m)$ can be obtained as

$$\overline{\{c(nT, m)\}^2} = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \{c(nT, m)\}^2 = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} |G^*(e^{j\omega T}, m)|^2 S_r^*(\omega) d\omega \quad \ldots\ldots [20]$$

Similarly, the mean-square value of the output samples $c(nT)$ at the sampling instants can be obtained from Equations [13] and [19].

Evidently, the relations of Equation [18] to [20] hold for a closed-loop feedback system as well as for an open-loop system. For instance, substituting the transfer function of Equation [20] by over-all transfer function of a feedback system shown in Fig. 4, we have

$$\overline{\{c(nT, m)\}^2} = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left| \frac{G^*(e^{j\omega T}, m)}{1 + G^*(e^{j\omega T})} \right|^2 S_r^*(\omega) d\omega \quad [21]$$

Similarly, for the error samples $\epsilon(nT)$

$$\overline{\{\epsilon(nT)\}^2} = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \left| \frac{1}{1 + G^*(e^{j\omega T})} \right|^2 S_r^*(\omega) d\omega \quad \ldots [22]$$

It is possible to design a minimum mean-square system by deriving an integral equation like the Wiener-Hopf equation and solving (11) to find the physically realizable form of $G^*(e^{j\omega T})$ to minimize $\overline{\{\epsilon(nT)\}^2}$. The realization of the best form of $G^*(e^{j\omega T})$ can be done easily by discrete compensation (4, 5).

It should be noted that by these relations we can evaluate the mean-square value of the output only with respect to specified discrete instants $(n - \Delta)T$. A method to obtain mean-square value with respect to over-all time t in continuous sense will be proposed in the next section.

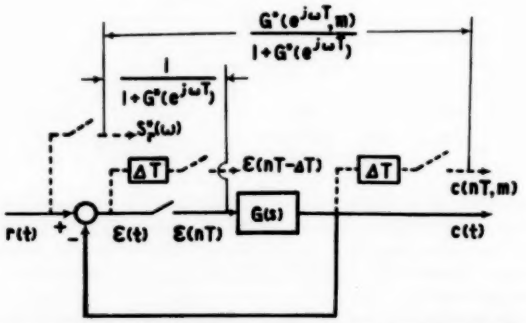


FIG. 4 A SAMPLED-DATA CONTROL SYSTEM. $r(t)$: REFERENCE INPUT; $c(t)$: CONTROLLED VARIABLE; $c(nT, m)$: SAMPLES OF CONTROLLED VARIABLE AT ANY INSTANTS; $\epsilon(t)$: ERROR; $\epsilon(nT)$: ERROR SAMPLES AT SAMPLING INSTANTS; $\epsilon(nT - \Delta T)$: ERROR SAMPLES AT ANY INSTANTS

Furthermore, the author has found also the useful relations for determination of the modified pulse-transfer function. In Fig. 2, if $S_{re}^*(\omega, m)$ is cross-spectral density of input samples at the sampling instants $r(nT)$ and output samples at any time $c(n - 1 + m)T$, the following equation can be obtained

$$S_{re}^*(\omega, m) = G^*(e^{j\omega T}, m)S_r^*(\omega) \quad [23]$$

The relation is reduced to the simpler form

$$S_{re}^*(\omega) = G^*(e^{j\omega T})S_r^*(\omega) \quad [24]$$

for (a special case) the pulse-transfer function $G^*(e^{j\omega T})$ instead of $G^*(e^{j\omega T}, m)$. The detailed derivations of Equations [23] and [24] are given in Appendix 2.

Since $S_r^*(\omega)$ is a real quantity, the phase of $S_{re}^*(\omega, m)$ is the same as that of $G^*(e^{j\omega T}, m)$. Equations [23] and [24] give novel means of determining the pulse-transfer functions of linear systems. This is analogous to the continuous case. This method has two significant advantages: (*a*) One can determine the dynamic characteristics of industrial process from normal operating record without upsetting the system (10). (*b*) Any disturbance other than the random input used for measurement does not appreciably affect the result, if there exists no cross correlation between the disturbance and the input.

If the input is a random signal distributed uniformly over the frequency range from zero to one half of the sampling frequency $f_0/2$, there exists the following relation as shown in Bibliography (6)

$$S_r^*(\omega) = PT = \text{const} \quad [25]$$

where

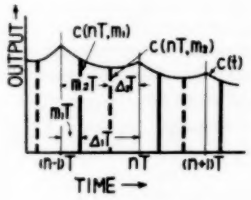
$$P = \overline{\{r(nT)\}^2} \quad [26]$$

Applying Equation [25] to Equation [23], we get the following simpler form which is a powerful relation for determining the system transfer function

$$G^*(e^{j\omega T}, m) = S_{re}^*(\omega, m)/PT \quad [27]$$

Mean Square of Absolute Value of Modified Pulse-Frequency Transfer Function

Generally, the output $c(t)$ of a sampled-data system is not limited to discrete sampled values. Thus the information required in many cases is mean square of $c(t)$ rather than that of sampled values $c(nT)$. Use of mean square of the modified pulse-frequency transfer function which is proposed in this section makes it possible to calculate the mean-square value of the total output $c(t)$.



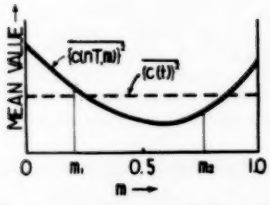


Fig. 5 Illustration for Equation [28]

From Equation [20], for each value of m, we get respective mean-square value of output samples as shown in Fig. 5. Therefore Equation [20] must be averaged from $m = 0$ to $m = 1$ to obtain the mean-square value of the total output $c(t)$

$$\overline{\{c(t)\}^2} = \frac{1}{1-0}\int_0^1 \left[\frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} |G^*(e^{j\omega T}, m)|^2 S_r^*(\omega)d\omega\right] dm \quad [28]$$

Inversion of the order of integration, which is allowable in this case, gives

$$\overline{\{c(t)\}^2} = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T}\left[\int_0^1 |G^*(e^{j\omega T}, m)|^2 dm\right] S_r^*(\omega)d\omega \quad [29]$$

Let us designate the mean square, with respect to m, of absolute value of modified pulse-frequency transfer function by $M(\omega)$

$$M(\omega) = \frac{1}{1-0}\int_0^1 |G^*(e^{j\omega T}, m)|^2\, dm = \overline{|G^*(e^{j\omega T}, m)|^2} \quad [30]$$

Then Equation [29] can be rewritten as

$$\overline{\{c(t)\}^2} = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} M(\omega)\, S_r^*(\omega)d\omega \quad [31]$$

It is not a very complicated procedure to derive the $M(\omega)$ from published tables of modified z-transforms (1, 3). As one can see in the tables (or in Table 1), m is always involved only in numerators of the modified pulse-transfer functions, in the forms of m^k, $e^{-am}T$, $\sin amT$, or $\cos amT$. Thus we can calculate $M(\omega)$ easily with the help of the tables of the modified z-transforms.

We can obtain mean-square value of the total controlled variable of the sampled-data sytem. For instance, in the system in Fig. 4

$$M(\omega) = \overline{\left|\frac{G^*(e^{j\omega T}, m)}{1 + G^*(e^{j\omega T})}\right|^2} = \frac{\overline{|G^*(e^{j\omega T}, m)|^2}}{|1 + G^*(e^{j\omega T})|^2} \quad [32]$$

Thus

$$\overline{\{c(t)\}^2} = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} \overline{\left|\frac{G^*(e^{j\omega T}, m)}{1 + G^*(e^{j\omega T})}\right|^2} S_r^*(\omega)d\omega \quad [33]$$

It should be noted that the $\overline{\{c(t)\}^2}$ calculation given here does not directly lead to the evaluation of mean-square error $\overline{\{\epsilon(t)\}^2}$ of the system, because the error samples $\epsilon(nT - \Delta T)$ at any instants cannot be represented by its input samples $r(nT)$ at the sampling instants and by the modified pulse-transfer function. Hence the importance of variable-system theory as developed in the following section, from which we can get the following end formula

$$\overline{\{\epsilon(t)\}^2} = \overline{\{r(t)\}^2} - \frac{1}{\pi}\int_{-\infty}^{\infty} \text{Re}\left[\frac{G(j\omega)}{T\{1 + G^*(e^{j\omega T})\}}\right] S_r(\omega)d\omega + \overline{\{c(t)\}^2} \quad [34]$$

where Re is the symbol for the real part and $S_r(\omega)$ is spectral density of continuous reference input.

Formulas on the Basis of Variable-System Theory

In this section, we present a method to obtain mean-square value of the total error of a typical system shown in Fig. 7 for which the method in the foregoing section fails.

The basic relation for the present purpose has been derived by Zadeh (12, 13), regarding the sampled-data system as a linear varying-parameter system.[4] His formula is

$$\overline{\{c(t)\}^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \overline{|K(j\omega, t)|^2}\, S_r(\omega)d\omega \quad [35]$$

where $S_r(\omega)$ is spectral density of the continuous input. Here the $K(j\omega, t)$ is termed as system function and is defined as the ratio of output $c(t)$ and input when the input is $e^{j\omega t}$. It contains t as a parameter

$$K(j\omega, t) = c(t)/e^{j\omega t} \quad [36]$$

Let us first look at the simpler system shown in Fig. 6, for which the output $\epsilon(t)$ for input $e^{j\omega t}$ can be obtained by Linvill's method (14) in the following form

$$\epsilon(t) = e^{j\omega t} - \frac{1}{T}\sum_{n=-\infty}^{\infty} G(j\omega + jn\omega_0)e^{(j\omega+jn\omega_0)t} \quad [37]$$

Substituting $c(t)$ in Equation [36] by $\epsilon(t)$ of Equation [37], and taking the mean square of $K(j\omega, t)$, we have

$$\overline{|K(j\omega, t)|^2} = 1 - \frac{2}{T}\,\mathrm{Re}[G(j\omega)] + \frac{1}{T^2}\sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 \quad [38]$$

where Re is the symbol for the real part. The mean values of the products between different frequencies are zero, hence the simpler result.

The author suggests the following form (see Appendix 3 for the proof) for the evaluation of last right-hand side terms of Equation [38]

$$\frac{1}{T^2}\sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 = \overline{|G^*(e^{j\omega T}, m)|^2} = M(\omega) \quad [39]$$

[4] This is abbreviated as variable system or variable network generally.

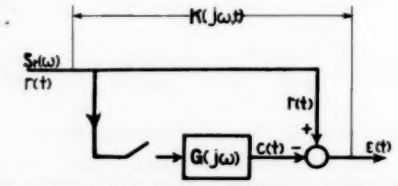


Fig. 6 System Which Requires Variable-System Theory

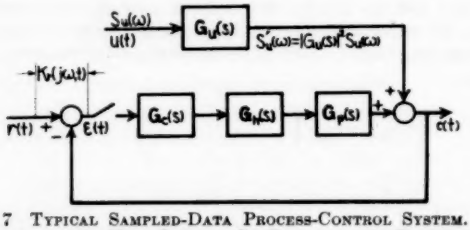


Fig. 7 Typical Sampled-Data Process-Control System. G_c: Controller; G_h: Hold Circuit; G_p: Process; $u(t)$: Disturbance; $S'_u(\omega)$: Continuous Spectral Density of Equivalent Disturbance

which he considers easier for mathematical processing than the following alternate form derived by Sklansky and Ragazzini (15)

$$\sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 = T\mathbf{Z}[G(s)G(-s)]_{z=e^{j\omega T}} \quad [40]$$

where $\mathbf{Z}[G(s)G(-s)]$ is the z-transform of the impulsive response of the system whose transfer function is $G(s)G(-s)$.

The same procedure applies to the control system given in Fig. 7. The system function for reference input is

$$K_r(j\omega, t) = \frac{\epsilon(t)}{e^{j\omega t}} = 1 - \frac{\sum_{n=-\infty}^{\infty} G(j\omega + jn\omega_0)e^{jn\omega t}}{T\{1 + G^*(e^{j\omega T})\}} \quad [41]$$

where $G^*(e^{j\omega T})$ is pulse-frequency-transfer function of the system whose continuous transfer function is $G_c(s)G_h(s)G_p(s)$.

Therefore

$$\overline{|K_r(j\omega, t)|^2} = 1 - 2\mathrm{Re}\left[\frac{G(j\omega)}{T\{1 + G^*(e^{j\omega T})\}}\right] + \frac{\sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2}{T^2|1 + G^*(e^{j\omega T})|^2} \quad [42]$$

Applying the relation of Equation [39] to the last term of Equation [42], we have the final formula

$$\overline{\{\epsilon(t)\}^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left\{1 - 2\mathrm{Re}\left[\frac{G(j\omega)}{T\{1 + G^*(e^{j\omega T})\}}\right] + \overline{\left|\frac{G^*(e^{j\omega T}, m)}{1 + G^*(e^{j\omega T})}\right|^2}\right\} S_r(\omega)d\omega \quad [43]$$

The same end formula applies for disturbance input, only replacing $S_r(\omega)$ in Equation [43] by $S_u{}'(\omega)$ which is shown in Fig. 7.

One of the future problems might be the analytical evaluation of the integral in Equation [43], which will be necessary for minimum mean-square-system synthesis.

Illustrative Examples

Example 1. To illustrate the procedures presented, the performance of the two control systems shown in Figs. 8(*a* and *b*) will be compared assuming the same process transfer function $1/(s + 1)$ and equal sampling period T for both systems. Both controllers consist of the discrete pulsed circuit such as for digital computers, but have different pulse-transfer functions. At system (*a*), pulse-transfer function of the controller $G_{ca}{}^*(z)$ is adjusted in such a way as to give the following pulse-transfer function from the reference input to the error

$$W^*(z) = 1 - z^{-1} \quad [44]$$

Hence

$$|W^*(e^{j\omega T})|^2 = 2 - 2\cos\omega T \quad [45]$$

From Equation [44] it is clear that the indicial response ends with finite settling time T as illustrated in Fig. 9(*a*).

At system (*b*), $G_{cb}{}^*(z)$ of the controller is adjusted so that $W^*(z)$ takes the form

$$W^*(z) = 1 - 2z^{-1} + z^{-2} \quad [46]$$

Then

$$|W^*(e^{j\omega T})|^2 = 6 - 8\cos\omega T + 2\cos 2\omega T \quad [47]$$

The indicial response of the system is illustrated in Fig. 9(*b*).

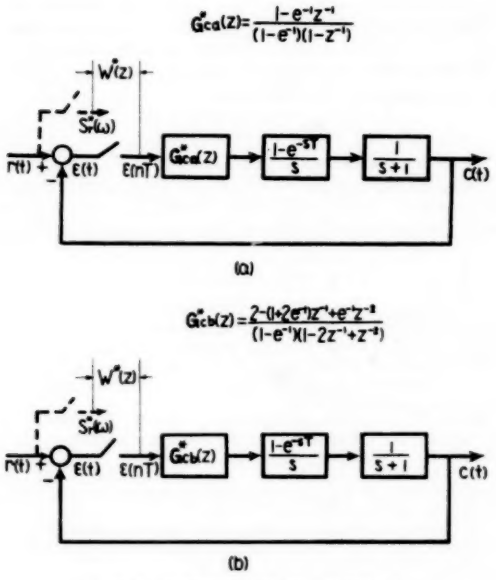


FIG. 8 SYSTEMS USED IN EXAMPLE 1

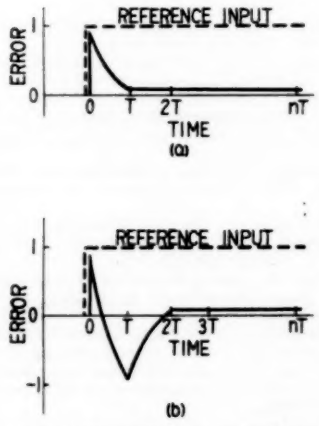


FIG. 9 INDICIAL RESPONSES OF THE SYSTEMS IN FIG. 8

Evidently the indicial response of system (a) is better than that of system (b), but it is not true for the random input. By Equation [20] or [22]

$$\overline{\{\epsilon(nT)\}^2} = \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} |W^*(e^{j\omega T})|^2\, S_r^*(\omega)\, d\omega \quad [48]$$

System (b) has a smaller value of $|W^*(e^{j\omega T})|^2$ than system (a) for lower value of ωT than $\pi/3$ as shown in Fig. 10. Therefore, if the fundamental frequency[4] distribution of $S_r^*(\omega)$ is limited to the lower frequency side of point P in the figure, which is the case in actual control systems (12), the system (b) has better performance than system (a). The statements here are based only on sampling instants.

Example 2. To illustrate the mean-square method on the

[4] It is enough to consider in the lowest range of frequency, since $S_r^*(\omega)$ is a periodic function with period f_0.

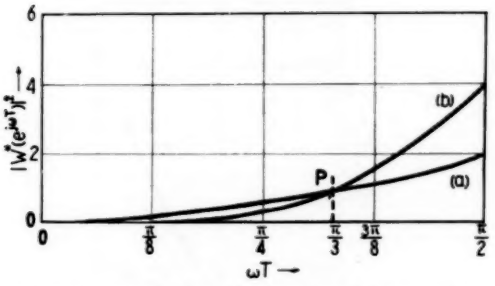


FIG. 10 PLOTS OF $|W^*(e^{j\omega T})|^2$ IN EXAMPLE 1 FOR ωT

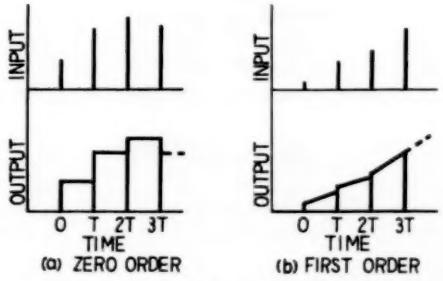


FIG. 11 TWO TYPES OF HOLD CIRCUIT

basis of variable-system theory taking the system of Fig. 6, two types of holding will be compared. If G in Fig. 6 is the hold circuit, the smaller $\overline{\{\epsilon(t)\}^2}$ may be regarded as indicating better performance—zero for ideal hold circuit which exactly reproduces the original continuous input.

For G of zero order hold shown in Fig. 11(a)

$$G(s) = \frac{1 - e^{-sT}}{s} \quad [49]$$

and

$$G^*(e^{j\omega T}, m) = e^{-j\omega T} \quad [50]$$

Therefore, from Equation [39]

$$M(\omega) = 1 \quad [51]$$

From Equation [49], for $T = 1$

$$\mathrm{Re}[G(j\omega)] = \frac{\sin\omega}{\omega} \quad [52]$$

Applying Equations [51] and [52] to [38]

$$\overline{|K(j\omega, t)|^2} = 2\left(1 - \frac{\sin\omega}{\omega}\right) \quad [53]$$

This is plotted as curve (a) in Fig. 12.

If G is the first-order hold as shown in Fig. 11(b)

$$G(s) = (1 - e^{-sT})^2\left(\frac{1}{s} + \frac{1}{Ts^2}\right) \quad [54]$$

and

$$G^*(e^{j\omega T}, m) = \frac{m(e^{j\omega T} - 1) + e^{j\omega T}}{e^{j2\omega T}} \quad [55]$$

Assuming $T = 1$, the third term of Equation [38] can be obtained easily by Equation [39]

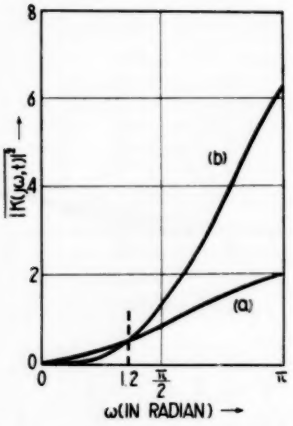


FIG. 12 CURVES OF THE SYSTEM FUNCTIONS USED IN EXAMPLE 2

$$M(\omega) = \frac{8 - 5\cos\omega}{3} \quad \ldots\ldots\ldots\ldots\ldots [56]$$

One can see Equation [40] is more laborious for the purpose than Equation [39]. The second term of Equation [38] becomes

$$\mathrm{Re}[G(j\omega)] = \frac{1}{\omega}(2\sin\omega - \sin 2\omega) - \frac{1}{\omega^2}(1 - 2\cos\omega + \cos 2\omega) \ldots\ldots [57]$$

Thus

$$\overline{|K(j\omega, t)|^2} = 1 - 2\left\{\frac{1}{\omega}(2\sin\omega - \sin 2\omega) - \frac{1}{\omega^2}(1 - 2\cos\omega + \cos 2\omega)\right\} + \frac{1}{3}(8 - 5\cos\omega) \ldots\ldots [58]$$

This is plotted as curve (*b*) in Fig. 12.

Therefore, by Equation [35], it may be concluded that the first-order hold is better than the zero-order hold when the spectral density of the input is distributed below $\omega = 1.2$. When the input is a white noise distributed from $\omega = 0$ to $\omega = \pi$, the former is worse than the latter.

CONCLUSION

The paper gives basic mathematical relations for the statistical treatments of sampled-data systems. The relations are analogous to those for continuous-data systems as summarized in Table 2.

The mean square of absolute value of the modified pulse-frequency-transfer function plays an important role not only for evaluation of mean-square value of its total output waveform, but also for the calculation on the basis of variable-system theory.

The importance of statistical approach of the sampled-data system is demonstrated in the examples.

ACKNOWLEDGMENTS

The author is sincerely grateful to Prof. Y. Takahashi for contributing his time and effort in guiding this research. Thanks are also expressed to Dr. M. Terao, Mr. E. Kikuchi, and Mr. T. Mitsumaki for their associations and constructive criticisms.

BIBLIOGRAPHY

1 "Frequency Response (ASME)," by R. Oldenburger, The Macmillan Company, New York, N. Y., 1956. See paper by R. H. Barker, or "The Pulse Transfer Function and Its Application to Sampling Servo Systems," by R. H. Barker, Proceedings of the Institution of Electrical Engineers, part IV, vol. 99, Monograph No. 43, July, 1952, pp. 302–317.

2 "Theory of Servomechanisms," by H. M. James, N. B. Nichols, and R. S. Phillips, McGraw-Hill Book Company, Inc., New York, N. Y., 1947, Chapter 6.

3 "Synthesis and Critical Study of Sampled-Data Control Systems," by E. I. Jury, Trans. AIEE (Applications and Industry), no. 25, vol. 75, July, 1956, pp. 141–151.

4 "Sampled-Data Processing Techniques for Feedback Control Systems," by A. R. Bergen and J. R Ragazzini, Trans. AIEE, part II, vol. 73, November, 1954, pp. 236–247.

5 "Discrete Compensation of Sampled-Data and Continuous Control Systems," by E. I. Jury and W. Schroeder, Electronics Research Laboratory Report, no. 154, University of California, Series 60, 1955.

6 "Information Theory," by S. Goldman, Prentice-Hall, Inc., New York, N. Y., 1953, chapters 2 and 8.

7 "Frequency Analysis of Digital Computers Operating in Real Time," by J. M. Salzer, Proceedings of the IRE, vol. 42, February, 1954, pp. 457–466.

8 "The Analysis of Sampled-Data Systems," by J. R. Ragazzini and L. A. Zadeh, Trans. AIEE, part II, vol. 71, November, 1952, pp. 225–234.

9 "Analysis and Synthesis of Sampled-Data Control Systems," by E. I. Jury, Trans. AIEE, part I, vol. 73, September, 1954, pp. 332–346.

10 "Determination of System Characteristics from Normal Operating Records," by T. P. Goodman and J. B. Reswick, Trans. ASME, vol. 78, 1956, pp. 256–271.

11 "Random Processes in Automatic Control," by J. H. Laning, Jr. and R. H. Battin, McGraw-Hill Book Company, Inc., New York, N. Y., 1956, chapter 7.

12 "Frequency Analysis of Variable Networks," by L. A. Zadeh, Proceedings of the IRE, vol. 38, March, 1950, pp. 291–299.

13 "Correlation Functions and Power Spectra in Variable Networks," by L. A. Zadeh, Proceedings of the IRE, vol. 38, November, 1950, pp. 1342–1345.

14 "Sampled-Data Control Systems Studied Through Comparison of Sampling with Amplitude Modulations," by W. K. Linvill, Trans. AIEE, vol. 70, 1951, pp. 1779–1788.

15 "Analysis of Errors in Sampled-Data Feedback Systems," by J. Sklansky and J. R. Ragazzini, Trans. AIEE, part II, vol. 74, May, 1955, pp. 65–71.

TABLE 2 CORRESPONDENCE BETWEEN SAMPLED-DATA SYSTEM AND CONTINUOUS-DATA SYSTEM

	Sampled-data system	Continuous-data system
Correlation function	$R_r^*(k) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} r(nT)r\{(n+k)T\}$	$R_r(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} r(t)r(t+\tau)dt$
Spectral density	$S_r^*(\omega) = T \sum_{k=-\infty}^{\infty} R_r^*(k)e^{-j\omega kT}$	$S_r(\omega) = \int_{-\infty}^{\infty} R_r(\tau)e^{-j\omega\tau}d\tau$
	At sampling instants	
Mean square	$\overline{[r(nT)]^2} = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} S_r^*(\omega)d\omega$	$\overline{[r(t)]^2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_r(\omega)d\omega$
Output-input relations	$S_e^*(\omega, m) = \lvert G^*(e^{j\omega T}, m)\rvert^2 S_r^*(\omega)$	$S_e(\omega) = \lvert G(j\omega)\rvert^2 S_r(\omega)$
	$S_{re}^*(\omega, m) = G^*(e^{j\omega T}, m)\, S_r^*(\omega)$	$S_{re}(\omega) = G(j\omega)\, S_r(\omega)$
	At sampling instants	
	$\frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \lvert G^*(e^{j\omega T})\rvert^2 S_r^*(\omega)d\omega$	
	Total[a]	
Mean square of output	$\frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \overline{\lvert G^*(e^{j\omega T}, m)\rvert^2} S_r^*(\omega)d\omega$ ..... [a]	$\frac{1}{2\pi} \int_{-\infty}^{\infty} \lvert G(j\omega)\rvert^2 S_r(\omega)d\omega$
	$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\lvert K(j\omega, t)\rvert^2} S_r(\omega)d\omega$ ..... [b]	

[a] Formula [b] can be used for the system to which formula [a] is not applicable.

Appendix 1

Correspondence Between Sampled-Data System and Continuous-Data System

Relations mentioned in the present paper are analogous to those for the continuous-data systems. Table 2 shows this analogy for reference.

Appendix 2

Derivations of Output-Input Relations

Derivation of Equation [18]. Any stable linear sampled-data system can be represented by a modified weighting sequence (2) on the past of the input samples $r(nT)$. Let the modified weighting sequence be written as $g(hT - \Delta T)$, where h is an integer.

Then $g(hT - \Delta T) = 0$ for $h < 0$

and $\sum_{h=0}^{\infty} \lvert g(hT - \Delta T)\rvert < \infty$ ..... [59]

Then the output sample $c(nT - \Delta T)$ is

$$c(nT - \Delta T) = \sum_{h=0}^{\infty} r(nT - hT)g(hT - \Delta T) \ldots [60]$$

The autocorrelation function of the output can be written in terms of the input as follows

$$R_e^*(k, \Delta) = \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \left[\sum_{h=0}^{\infty} r(nT - hT)g(hT - \Delta T) \times \sum_{q=0}^{\infty} r(nT + kT - qT)g(qT - \Delta T)\right]$$

$$= \lim_{N\to\infty} \frac{1}{2N+1} \sum_{h=0}^{\infty} \sum_{q=0}^{\infty} \left[\sum_{n=-N}^{N} r(nT - hT)r(nT + kT - qT)\right] g(hT - \Delta T)g(qT - \Delta T)$$

$$= \sum_{h=0}^{\infty} \sum_{q=0}^{\infty} R_e^*(k + h - q)g(hT - \Delta T)g(qT - \Delta T) \ldots [61]$$

This is the equivalent form of Equation [18] in the time domain.

Equation [18] can now be proved by transforming this into the frequency domain with the relation of Equation [17]. Thus

$$S_e^*(\omega, \Delta) = T \sum_{k=-\infty}^{\infty} R_e^*(k, \Delta)e^{-j\omega kT}$$

$$= T \sum_{k=-\infty}^{\infty} \sum_{h=0}^{\infty} \sum_{q=0}^{\infty} [R_r^*(k + h - q)e^{-j\omega(k+h-q)T} \times g(hT - \Delta T)e^{j\omega hT}g(qT - \Delta T)e^{-j\omega qT}] \ldots [62]$$

Here the variable $(k + h - q)$ can be changed to k. Then

$$S_e^*(\omega, \Delta) = T \sum_{h=0}^{\infty} g(hT - \Delta T)e^{j\omega hT} \times \sum_{q=0}^{\infty} g(qT - \Delta T)e^{-j\omega qT} \sum_{k=-\infty}^{\infty} R_r^*(k)e^{-j\omega kT} \ldots [63]$$

Therefore, let $m = 1 - \Delta$

$$S_e^*(\omega, m) = G^*(e^{j\omega T}, m)G^*(e^{-j\omega T}, m)S_r^*(\omega) \ldots [64]$$

Thus we have gotten Equation [18].

Equation [19] can be obtained similarly.

Derivation of Equation [23]. Substituting Equation [60] in [8], the cross correlation function between input and output becomes

$$R_{re}^*(k, \Delta) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \left[r(nT) \sum_{h=0}^{\infty} r(nT + kT - hT) g(hT - \Delta T) \right]$$

$$= \sum_{h=0}^{\infty} \left[g(hT - \Delta T) \left\{ \lim_{N \to \infty} \frac{1}{2N+1} \sum_{n=-N}^{N} r(nT) r(nT + kT - hT) \right\} \right]$$

$$= \sum_{h=0}^{\infty} g(hT - \Delta T)\, R_r^*(k - h) \ldots\ldots [65]$$

We can get Equation [23] by transforming Equation [65] into the frequency domain by the same relation to Equation [17] as follows

$$S_{re}^*(\omega, \Delta) = T \sum_{k=-\infty}^{\infty} \left[\sum_{h=0}^{\infty} g(hT - \Delta T) R_r^*(k - h) \right] e^{-j\omega kT}$$

$$= T \left[\sum_{k=-\infty}^{\infty} R_r^*(k - h) e^{-j\omega(k-h)T} \right] \times \left[\sum_{h=0}^{\infty} g(hT - \Delta T) e^{-j\omega hT} \right] \ldots\ldots [66]$$

Therefore, let $m = 1 - \Delta$

$$S_{re}^*(\omega, m) = G^*(e^{j\omega T}, m) S_r^*(\omega) \ldots\ldots [67]$$

Thus Equation [23] has been proved. Equation [24] can be derived in the same way.

Appendix 3

PROOF OF EQUATION [39]

With reference to Fig. 2, the modified pulse-transfer function of a system whose impulsive response is $g(t)$, or the modified z-transform of a wave form $g(t)$ is given as (3)

$$G^*(z, \Delta) = \mathfrak{L}[g(t - \Delta T)\delta_T(t)] = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} G(p) e^{-\Delta pT} \frac{1}{1 - e^{pT} z^{-1}}\, dp \ldots\ldots [68]$$

where $z = e^{sT}$, T = sampling period, $\delta_T(t)$ = a train of unit impulses, $\mathfrak{L}$ is symbol for the Laplace transformation $G(p) = \mathfrak{L}[g(t)]_{s=p}$, or let $m = 1 - \Delta$, $0 \leq \Delta < 1$, then

$$G^*(z, m) = e^{-smT} \mathfrak{L}[g(t + mT)\delta_T(t)] \ldots\ldots [69]$$

Evaluating the integral of Equation [68] in the closed contour formed by the line $c - j\infty$ to $c + j\infty$ and the infinite semi-circle which encloses the singularities of $1/(1 - e^{pT}z^{-1})$ in the right half p-plane, we can obtain Equation [4]

$$G^*(z, m) = \frac{1}{T} \sum_{n=-\infty}^{\infty} G(s + jn\omega_0) e^{-(s+jn\omega_0)(1-m)T} \ldots [70]$$

where $\omega_0 = 2\pi/T$ and n is an integer.

The square of $|G^*(z, m)|$ is

$$|G^*(e^{j\omega T}, m)|^2 = \frac{1}{T^2} \left[\sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 + \sum_{n=-\infty}^{\infty} \sum_{\substack{q=-\infty \\ n \neq q}}^{\infty} G(j\omega + jn\omega_0) G(-j\omega - jq\omega_0) e^{j(q-n)\omega_0(1-m)T} \right] \ldots\ldots [71]$$

The mean of the second summation is zero. Hence averaging Equation [71] from $m = 0$ to $m = 1$, we obtain

$$\overline{|G^*(e^{j\omega T}, m)|^2} = \frac{1}{T^2} \sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 \ldots\ldots [72]$$

Thus Equation [39] has been proved.

Discussion

G. F. FRANKLIN.[6] This is an interesting paper which presents some new formulas for the analysis of random inputs in sampled-data systems. It is the purpose of this discussion to point out several statements in the paper which the discusser feels should be clarified and to indicate an alternative derivation of some of the results of the paper.

The argument (a) of the author in the introduction would seem to have limited validity because of the existence of design procedures which result in satisfactory responses to step and ramp inputs with no restrictions caused by a relation between the sampling rate and the frequency content of the input (16, 17, 18).[7] Also, the statement in the paper following Equation [22] should be amplified. In 1955, two derivations of the optimum linear filter for the least squares prediction and smoothing of sampled-data were published (19, 20). Both of these papers considered the construction of continuous functions from discrete data.

A constant pulse spectral density as given by Equation [25] of the paper may be generated by a far wider class of inputs than is implied by the sentence preceding this equation. As a matter of fact, from Equations [7] and [17] it follows immediately that $S_r^*(\omega)$ will be a constant if successive samples $r(nT)$ are uncorrelated. The samples may or may not be independent, of course, but the autocorrelation function $\phi_r(\tau)$ must be zero at $\tau = kT$ for $k \neq 0$ to give the desired result.

The principal objective of the discussion is to present an alternative to the author's Equations [38] and [35] for the calculation of the total mean-square error of the sampled-data system. From Fig. 6 the mean-square error is given by

$$\bar{e}^2 = \bar{r}^2 - 2\overline{rc} + \bar{c}^2 \ldots\ldots [73]$$

The value of [73] can be calculated by the integral of the spectral densities of the three terms given in Equation [73]. These spectral densities are

$$S_r(\omega) = S_r(\omega) \ldots\ldots [74]$$

$$S_{rc}(\omega) = \frac{1}{T} S_r(\omega)\, G(j\omega) \ldots\ldots [75]$$

$$S_c(\omega) = \frac{1}{T^2} S_r^*(\omega)\, |G(j\omega)|^2 \ldots\ldots [76]$$

where $S_r^*(\omega)$ is as defined in the paper. The transfer function $G(j\omega)$ in Equations [75] and [76] is entirely general and for the feedback system of Fig. 4 of the paper would be replaced by

[6] Assistant Professor, Electrical Engineering Department, Columbia University, New York, N. Y.

[7] Numbers in parentheses refer to the Bibliography at the end of this discussion.

$$\frac{G(j\omega)}{1 + G^*(e^{j\omega T})}$$

The derivations of Equations [75] and [76] follow immediately from the input-output relation of the system and Equations [7] and [17] of the paper. The mean square error of the system is given by

$$\overline{e^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left\{ S_r(\omega) - \frac{2}{T} S_r(\omega) G(j\omega) + \frac{1}{T^2} S_r^*(\omega) |G(j\omega)|^2 \right\} d\omega \quad \ldots [77]$$

which may be shown to be equivalent to Equation [38] and [35] of the paper. The evaluation of mean square error by use of Equation [77] requires only one integration instead of the two suggested by the author and, in addition, relates this problem to the familiar stationary filter problem rather than to the variable network theory. As an example of the application of Equation [77], let the input be described by

$$S_r(\omega) = \frac{2\omega_0}{\omega^2 + \omega_0^2}$$

$$S_r^*(\omega) = \frac{T \sinh \omega_0 T}{\cosh \omega_0 T - \cos \omega t} \quad \ldots [78]$$

and let $G(j\omega)$ be the transfer function of a zero order hold given by Equation [49] of the paper. Then Equation [77] may be evaluated to give

$$\overline{e^2} = 2\left\{1 - \frac{1 - e^{-\omega_0 T}}{\omega_0 T}\right\} \quad \ldots [79]$$

which may be used to set the sampling period for a given mean square error.

BIBLIOGRAPHY

16 (1) of the paper.
17 (4) of the paper.
18 "Factors in the Design of Digital Controllers for Sampled Data Control Systems," by J. E. Bertram, Trans. AIEE, vol. 75, part II, 1956, p. 151.
19 "Linear Filtering of Sampled Data," by G. F. Franklin, 1955 IRE Convention Record, part IV, p. 119.
20 "Linear Least Squares Filtering and Prediction of Sampled Signals," by S. P. Lloyd and B. McMillan, Proceedings of the Symposium on Modern Network Synthesis, Polytechnic Institute of Brooklyn, Brooklyn, N. Y., 1955, p. 221.

E. I. Jury.[8] The author is to be commended for his extension of the modified Z-transform theory to the statistical analysis of sampled-data systems (21, 22, 23).[9] Table 2 of the paper illustrates clearly the various relations obtained for sampled data and its connection with the continuous theory. It serves a clear need in the analysis problem associated with error-sampled-data systems.

It is true as the author indicates that sampled-data control systems designed for deadbeat response for aperiodic inputs might behave adversely for random inputs. However, there exist many applications of sampled-data control systems where deadbeat response is very desirable and the utilization of digital computers to obtain such rigid performance is becoming increasingly important, especially in machine tool industry (24).

The various relations developed in this paper are also applicable to the analysis of control systems utilizing digital computers, for essentially such mixed systems can be treated as sampled-data systems with certain desired sampling period "T" which represents the duty cycle of the digital element.

A logical extension of the statistical analysis of sampled-data systems is the synthesis problem, whereby the discrete compensator parameters can be designed to fulfill certain index of performance. Investigations along these lines are presently being tackled at this institution and results obtained show promise of obtaining an optimum digital filter to tally with a certain criterion of design.

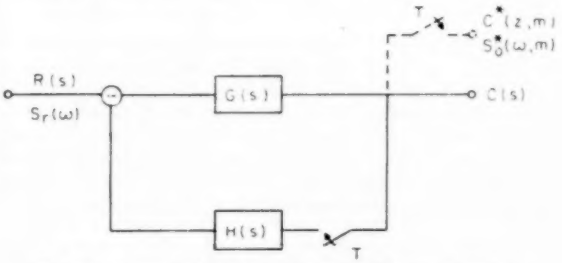


Fig. 13 Sampled-Data Control System Configuration

The statistical relations obtained by the author are mainly applicable to sampled-data control configurations shown in Figs. 6 and 7. However, there exist other types of configuration (25) which the discusser encountered in which neither the modified Z-transform approach nor the system function is readily applicable. For instance, consider the sampled-data control system shown in Fig. 13, the output modified Z-transform of which can be written as (26)

$$C^*(z, m) = RG^*(z, m) - \frac{RG^*(z)}{1 + HG^*(z)} HG^*(z, m) \quad \ldots [80]$$

It is noticed from Equation [80] that the input cannot be separated from the transfer function which complicates the process of obtaining the system function $K(s, t)$ to describe the input-output relationship. This type of difficulty is inherent in some sampled-data control system configurations which impose certain restrictions on the statistical analysis.

In view of the availability of extensive tables of Z-transforms (26) I found it easier in some cases to evaluate Equation [39] using relation [40] to avoid squaring and integration with respect to m. However, without such tables the author's contention in suggesting the evaluation of the last term of Equation [38] using relation [39] is principally correct.

Relation [27], developed by the author for determining the modified z-transform system function, is important and should be actually tested for actual systems.

It might be added that the analysis formulation obtained by Mr. Mori can also be applied for systems having pure delay (integer or non-integer of the sampling period) if the proper interpretation "m" is accordingly observed (27, 28).

In conclusion, the author's contributions in this paper are of considerable importance and the new method of statistical design will undoubtedly enhance the growing field of sampled-data control systems.

BIBLIOGRAPHY

21 "Interpolation and Extrapolation of Sampled Data," by A. B. Lees, IRE Trans. on Information Theory, 1956, pp. 12–17.
22 "Über die Synthese von Impulssystemen der automatischen Regelung und Steuerung," by J. Z. Cypkin, Fachtagung Regelungstechnik, Heidelberg, Germany, 1956, Beitrag Nr. 95 Unkorrigierter Vordruck mit 18 Bildern.
23 "An Extension of the Minimum-Square Error Theory for

[8] University of California, Berkeley, Calif.
[9] Numbers in parentheses refer to the Bibliography at the end of this discussion.

Sampled-Data," by M. Blum, Trans. IRE, vol. IT-2, no. 3, 1956, pp. 176–184.

24 "Sampled-Data Systems, by D. J. Gimbel, *Control Engineering*, vol. 4, no. 2, 1957.

25 "Table of Z-Transform and Modified Z-Transforms of Various Sampled-Data Control Systems Configurations," by E. I. Jury and G. Parmanfarma, I. E. R. Series 60, no. 136A, 1955, University of California, Berkeley, Calif.

26 (1) and (9) of the paper.

27 "Analysis and Synthesis of Sampled-Data and Continuous Control Systems With Pure Time Delays," by W. Schroeder, I. E. R. Series 60, no. 156, 1956, University of California, Berkeley, Calif.

28 "Additions to the Modified Z-Transform Theory," by E. I. Jury, submitted to IRE, 1957.

JACK SKLANSKY.[10] While the variable-system approach to evaluating the mean square error of sampled-data systems is important, there exists an interesting alternative approach involving an evaluation of the spectral density of the error and its integral over the real frequency axis. One uses a derivation similar to Stewart's (29)[11, 12], which is based on finding the limit of the "average" of the square of the Fourier transform of the truncated error as the truncation interval approaches infinity, viz.

$$S_\epsilon(\omega) = \lim_{\tau \to \infty} \frac{1}{\tau} \left| \mathfrak{F}\epsilon_\tau(t) \right|^2, \quad \ldots\ldots\ldots\ldots\ldots [81]$$

where $\epsilon_\tau(t)$ is equal to $\epsilon(t)$ inside a time interval τ units long, and is zero outside that interval. The result is

$$S_\epsilon(\omega) = S_r(\omega) - \frac{2}{T} S_r(\omega) \mathrm{Re}\left[\frac{G(j\omega)}{1 + G^*(e^{j\omega T})}\right] + \frac{1}{T^2} S_r^*(\omega) \left|\frac{G(j\omega)}{1 + G^*(e^{j\omega T})}\right|^2 \quad \ldots. [82]$$

Integrating this over the entire range of negative and positive values of real frequency obtains the mean square error.

It will be noted that the right member of Equation [82] is the same as the integrand in Equation [43] except for the last terms in each. For consistency, therefore, the integral over the real frequency axis of each of these terms [which, by the way, happens to yield $\overline{c^2(t)}$] must be the same. That this is the case is demonstrated as follows: Note that the starred functions in the last term of Equation [82] are periodic in ω with period ω_0; as a result of this periodicity, the integral of that term over the entire real frequency axis is equal to

$$\overline{c^2(t)} = \frac{1}{2\pi T} \int_{-\omega_0/2}^{\omega_0/2} \frac{S_r^*(\omega)}{|1 + G^*(e^{j\omega T})|^2} \sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 d\omega \ldots. [83]$$

Making the transformation $e^{j\omega T} \to z$ obtains

$$\overline{c^2(t)} = \frac{1}{2\pi j} \oint_{\substack{\text{unit} \\ \text{circle}}} \frac{S_r^*(z)\, \Phi_g^*(z)}{T[1 + G^*(z)]\,[1 + G^*(z^{-1})]} z^{-1}\, dz \ldots [84]$$

where $\Phi_g^*(z)$ is defined by

$$\Phi_g^*(z) = \left\{ \frac{1}{T} \sum_{n=-\infty}^{\infty} |G(j\omega + jn\omega_0)|^2 \right\}_{e^{j\omega T} = z} \quad \ldots. [85]$$

[10] David Sarnoff Research Center, Radio Corporation of America, Princeton, N. J.

[11] Numbers in parentheses refer to the Bibliography at the end of this discussion.

[12] Work of a related nature has also been published by Franklin (30) and Lloyd and McMillan (31).

$[\Phi_g^*(z)$ can be evaluated by either Equation [39] or [40], with the modification that the variable z should not be replaced by $e^{j\omega T}$.] In a similar manner the integral of the last term of the integrand of Equation [43] can be shown to be equal to the right member of [84], thus obtaining the sought demonstration that the infinite integral of [82] equals the right member of [43].

A valuable by-product of the foregoing analysis is Equation [84], since it happens to be a formula for evaluating $\overline{c^2(t)}$ analytically. This formula apparently is either new or little known, since past authors, including the discusser, have generally been unable to suggest how to evaluate integrals like [43] other than numerically. To evaluate [84], one may either (*a*) find the residues of its integrand inside the unit circle of the z-plane, or (*b*) substitute $(1 + y)/(1 - y)$ for z and use the table in the Appendix of James, Nichols, and Phillips (32). A third alternative would be to compute a special table for integrals of this type.

It is sometimes useful to distinguish between the "system error" and the "control error": The first is the difference between the desired output and the actual output; the second is the difference between the input and the fedback signal. In Mr. Mori's paper, the two terms can coincide in meaning; "control error," however, is a helpful additional term when a distinction in meaning is desired.

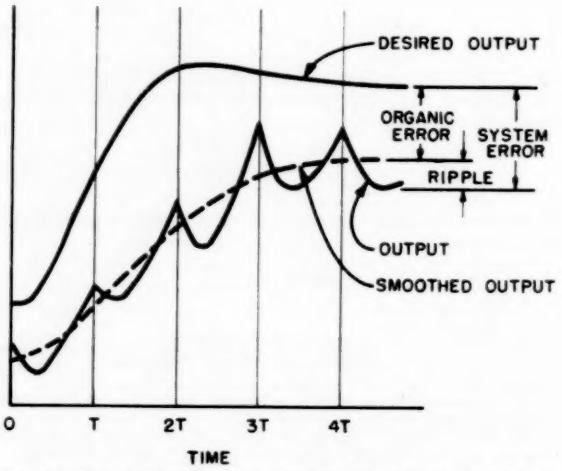


FIG. 14 ROUGH INTUITIVE DESCRIPTION IN THE TIME DOMAIN OF RIPPLE, ORGANIC ERROR, AND SYSTEM ERROR

In discussing the system error of sampled-data systems, it is often instructive to resolve the system error into that caused by the dynamic lags and leads and that caused by the sampling process. These component errors in the past have been referred to as "organic error" and "ripple," respectively (33); they are so denoted in Fig. 14, where they are indicated in a rough, intuitive manner in the time domain.

A useful mathematical phenomenon occurs when the system error and the control error are identical: The spectral densities of the ripple and organic error then sum to the spectral density of the system error, and, as a consequence, the mean square ripple and the mean square organic error sum to the mean square system error. This is shown by expressing Equation [82] in the form

$$S_\epsilon(\omega) = S_r(\omega) \left|1 - \frac{G(j\omega)}{T[1 + G^*(e^{j\omega T})]}\right|^2 + [S_r^*(\omega) - S_r(\omega)] \left|\frac{G(j\omega)}{T[1 + G^*(e^{j\omega T})]}\right|^2 \quad \ldots. [86]$$

The first term of the right member is the spectral density of the organic error, and the second term is the spectral density of the ripple.

The author's Equation [39], which provides a method alternative to that of Equation [40] for evaluating the infinite sum of $|G(j\omega + jn\omega_0)|^2$, is certainly valuable. However, the discusser questions the author's contention that [39] is easier to use than [40]. It seems to the discusser that the contention would only be valid if the usual tables of z-transforms give $G^*(z, m)$ and not $\mathbf{Z}[G(s)G(-s)]$. But this is not ordinarily the case. In fact, there are several good tables—such as Jury's (34), Stone's (35), and Truxal's (36)—which don't list $G^*(z, m)$ at all.

Even when (z, m)-transforms are tabulated, the discusser cannot find a distinct advantage of one formula over the other, except in a certain case to be mentioned shortly. Consider, for example, the author's Example 2. Using Equation [39], one first has to find the (z, m)-transforms of $1/s$ and $1/s^2$ from a table such as Barker's (37), add them, and replace z by $e^{j\omega T}$ to obtain Equation [55]. One then squares Equation [55] and averages with respect to m to obtain the desired result.

With the technique of Equation [40], one multiplies [54] by $G(-s)$ to obtain

$$G(s)G(-s) = (1 - e^{-sT})^2 (1 - e^{sT})^2 \left(-\frac{1}{s^2} + \frac{1}{T^2 s^4} \right) \ldots [87]$$

Both $1/s^2$ and $1/s^4$ are listed in Barker's table, and the evaluation of

$$\mathbf{Z}[G(s)G(-s)]_z = e^{j\omega T}$$

follows immediately. Thus in this example no clear difference in computational labor between the two techniques is apparent.

However, there is a certain case, of relatively infrequent occurrence, in which the tables favor $G^*(z, m)$ over $\mathbf{Z}[G(s)G(-s)]$, namely, the case where $G(s)$ has one or more multiple poles off the origin. Suppose $G(s)$ is expandable into partial fractions of the form $p_i/(s + a_i)^m$. Then $G(s)G(-s)$ can be expanded into fractions of the form $q_i/(s^2 - a_i^2)^m$. To each of the latter fractions, one can assign a z-transform; however, presently available tables (among which Barker's seems to be the most extensive) don't list these z-transforms for the cases where $m \geq 2$, while Barker's table does give the z-transforms of $p_i/(s + a_i)^m$ for all positive integral m. Thus when $G(s)$ has multiple poles, deficiencies in the available tables make Equation [39] easier to use than [40]. If the tables were extended to include the z-transform of $q_i/(s^2 - a_i^2)^m$ for $m \geq 2$, then the computational labor in Equations [39] and [40] for all rational forms of $G(s)$, including those with multiple poles off the origin, would seem to be about the same. Equation [40], by the way, could be used in making this extension.

Thus, as far as *inherent* computational labor is concerned, there is little difference that the discusser can see between Equations [39] and [40].

BIBLIOGRAPHY

29 "Statistical Design and Evaluation of Filters for the Restoration of Sampled Data," by R. M. Stewart, Proceedings of the IRE, vol. 44, no. 2, 1956, pp. 253–257.

30 "Linear Filtering of Sampled Data," by G. Franklin, 1955 IRE Convention Record, part 4, p. 119.

31 "Linear Least Squares Filtering and Prediction of Sampled Signals," by S. P. Lloyd and B. McMillan, Proceedings of the Symposium on Modern Network Synthesis, sponsored by the Polytechnic Institute of Brooklyn, Brooklyn, N. Y., 1955, p. 221.

32 (2) of the paper.

33 (15) of the paper.

34 (9) of the paper.

35 "A List of Generalized Laplace Transforms," by W. M. Stone, Iowa State College Journal of Science, vol. 22, no. 3, 1948, pp. 215–225.

36 "Automatic Feedback Control System Synthesis," by J. G. Truxal, McGraw-Hill Publishing Company, New York, N. Y., 1955.

37 "The Pulse Transfer Function and Its Application to Sampling Servo-Systems," by R. H. Barker, Proceedings of the IEE, part IV, monograph 43, July 15, 1952.

Author's Closure

The author greatly appreciates the interesting and constructive discussions of Professors Franklin, Jury, and Dr. Sklansky.

The author considers for Professor Franklin's statement about the argument (a) in the introduction, as follows: If inputs to sampled-data systems are limited to the step type, the frequency band limitation for interpolation or extrapolation is not necessary. However, the sampled-data control systems differ from the continuous-data control systems which show satisfactory responses for usual inputs if they show good responses for the step type inputs. In the sampled-data control systems, the excellent responses for the step type inputs form a special case as described in the argument (b) of the author in the introduction. Accordingly, it should be noted that the indicial response method which utilizes such a special input as step or ramp is liable to overestimate the controllabilities of the sampled-data control systems without the attention to this specialty of the sampled-data systems.

The author considers that the constant pulse spectral density as given by Equation [25] should be defined by Equations [7] and [17].

Professor Franklin's derivation of Equation [77], including Dr. Sklansky's Equation [82], is very interesting and valuable. In fact, the evaluation of mean square $\overline{[e(t)]^2}$ by use of Equation [77] does not require such an integration with respect to m as Equation [30]. But, in many sampled-data control systems the most laborious part in the evaluation of total mean square error is the calculation of the second term of the right-hand side of Equation [77] or [43]. In many cases, the second term in the integrand is a function of $j\omega$ and $e^{j\omega T}$, and the third term is a function only of $e^{j\omega T}$. Therefore the second term would be more complicated than the third term. A brief discussion which shows the existence of hidden response in the controlled variable of a sampled-data control system involving dead time would serve an example. Let us consider that the process of Fig. 8(a) has dead time $\Delta' T$ less than one sampling period T. And in the following $T = 1$ is assumed. Then the transfer function of the process becomes $e^{-s\Delta'}/(s + 1)$. For Equation [44] holds, the pulse transfer function of the controller $G_{ca}^*(z)$ should be given as

$$G_{ca}^*(z) = \frac{1 - e^{-1}z^{-1}}{(1 - e^{-m'}) - (1 - 2e^{-m'} + e^{-1})z^{-1} + (e^{-1} - e^{-m'})z^{-2}} \ldots [88]$$

where $m' = 1 - \Delta'$.

Applying Equation [42] to this system, the following equation can be obtained

$$\overline{|K(j\omega, t)|^2} = 1 - \frac{-2}{\omega(\omega^2 + 1)(A^2 + 2AB\cos\omega + B^2)} \times [\omega \{B\cos(-m')\omega + C\cos(1 - m')\omega - D\cos(2 - m')\omega + E\cos(3 - m')\omega\} + \{B\sin(-m')\omega + C\sin(1 - m')\omega - D\sin(2 - m')\omega + E\sin(3 - m')\omega\}] + \frac{(1 - 2e^{-1}\cos\omega + e^{-2}) - (1 - \cos\omega)(1 - e^{-2})}{A^2 + 2AB\cos\omega + B^2} \ldots [89]$$

where

$$A = 1 - e^{-m'}$$
$$B = e^{-m'} - e^{-1}$$
$$C = 1 + e^{-1} + e^{-2} - 2e^{-m'} - e^{-1-m'}$$

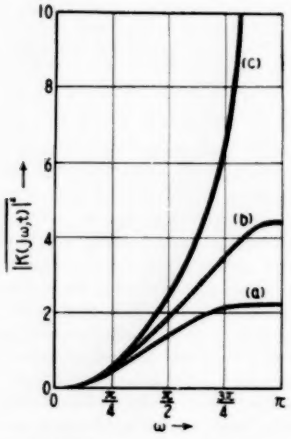


FIG. 15 PLOTS OF $\overline{|K(j\omega, t)|^2}$ FOR VARIOUS m'. (*a*) IS FOR $m' = 0$. (*b*) IS FOR $m' = 0.20$, (*c*) IS FOR $m' = 0.38$.

$$D = 1 + e^{-1} + e^{-2} - e^{-m'} - 2e^{-1-m'}$$
$$E = (1 - e^{-m'})e^{-1}$$

The second term of the right-hand side of Equation [89] is more complicated than the third term.

Equation [89] is plotted as shown in Fig. 15. In Fig. 15, curve (*a*) is for $m' = 0$, (*b*) is for $m' = 0.20$, (*c*) is for $m' = 0.38$. Because Equation [44] holds, for any value of m', the system has the same sampled error $\epsilon(nT)$ at the sampling instants, and consequently, it has same $\overline{\{\epsilon(nT)\}^2}$. But, one can see from Fig. 15 that the value of m', that is the value of the dead time, greatly affects the $\overline{|K(j\omega, t)|^2}$ of Equation [89]. Thus it can be concluded that, for the same random inputs involving considerable high frequency, the total mean square error $\overline{\{\epsilon(t)\}^2}$ is greatly affected by the value of the dead time though the sampled mean square error $\overline{\{\epsilon(nT)\}^2}$ is not affected by the value of the dead time.

The author agrees with Professor Jury on the points that the relations in the present paper are applicable to the analyses of control systems utilizing digital computers or having pure delay. The statistical treatment of the digital computer control system is one of the objectives of the present paper.

About the configurations of sampled-data systems pointed out by Professor Jury, the author considers as follows: The statistical relations obtained by the author are applicable also to such systems as shown in Fig. 13, after suitable system reductions. For instance, the system in Fig. 13 is transformable to the system as shown in Fig. 16(*a*) by adding new input and by moving the minus sign. And then the system can be reduced to system as shown in Fig. 16(*b*). It should be assumed that the newly added input signal has zero amplitude. We can find, if we regard the input $R(s)$ in Fig. 16(*b*) as the disturbance input in Fig. 7, that the system of

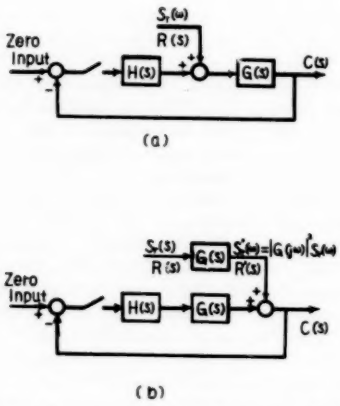


FIG. 16 SYSTEM REDUCTIONS OF THE SYSTEM OF FIG. 13 FOR APPLICATION OF EQUATION [43]

Fig. 16(*b*) is equivalent to that of Fig. 7. Therefore the system of Fig. 13 is equivalent to that of Fig. 7. Thus the author considers that Equation [43] which is applicable to the disturbance input of Fig. 7 is also applicable to the system of Fig. 13 by replacing $S_r(\omega)$ with $S_r'(\omega)$.

Dr. Sklansky's Equation [82] which is equal to the integrand of Professor Franklin's Equation [77] is valuable. It is very interesting that the following relation can be derived from Equation [43] and Equation [77] or [82]

$$\int_{-\infty}^{\infty} S^*(\omega)|G(j\omega)|^2 d\omega = \int_{-\infty}^{\infty} S(\omega)\overline{|G^*(e^{j\omega T}, m)|^2} d\omega \quad . . [90]$$

In the paper, only the unity feedback system is treated and consequently the system error coincides with the control error. The author considers that in a general sense the minimum system error might be desired for good control. The author agrees with Dr. Sklansky about the clear resolution of the system error into the organic error and the ripple which is raised by Dr. Sklansky. And he considers that the relation of Equation [86] is very interesting and useful.

About the difference between Equation [39] and Equation [40] on the labor of evaluation, the author considers as follows: Generally the evaluation of $Z[G(s)G(-s)]$ requires z-transformation of the function of twice orders compared with that of $G^*(z, m)$. Consequently, Equation [40] requires more extensive tables of z-transforms than Equation [39] on the assumption that the tables of modified z-transforms are provided. And for easy evaluation such general forms—$1/s^k$, $1/(s + a)^k$, $a/s^k(s + a)$, etc.—as found for higher order functions in Barker's table (1) or Jury's table (9) are unsatisfactory. In many cases, the evaluation of the higher order z-transforms by use of these general forms are more laborious than the evaluations of the squaring and of the integration with respect to m in Equation [39].

Optimization of Time-Varying Linear Systems With Nonstationary Inputs

By MARVIN SHINBROT,[1] MOFFETT FIELD, CALIF.

A method is presented for solving the integral equation which arises in optimization problems with nonstationary inputs. The method depends on the correlation functions being of a certain type—fortunately, a type which arises frequently in practice. The sort of problem which can be handled and the associated results are illustrated by examples.

Introduction

UNTIL the work of Wiener (1)[2] became known, the design of systems depended on a combination of cut-and-try procedures and analytical methods for choosing, in an optimum fashion, the free parameters of a system of given form. Wiener's principal contribution to design philosophy is the understanding that even the form as well as the parameters of a system can be chosen optimally.

Fundamental and important as the Wiener theory is, however, its scope is severely limited by the requirement that the inputs to the system be stationary. Thus, even so straightforward a problem as the question of the optimal design of a gun platform, say, which is to follow a target moving with constant speed in a straight line when the measurements are corrupted by noise, cannot be solved by Wiener's methods.

In an effort to eliminate this restriction, a new theory was devised in 1951 by Booton (2). By a method which is the direct generalization of the method used in (3), Booton derived the integral equation which a system with nonstationary inputs must satisfy in order to qualify as an optimum. However, he did not solve this equation.

Although to solve Booton's integral equation in full generality would be a fabulous accomplishment, it can be solved in certain circumstances. The problem then is to determine conditions sufficiently mild that many practical problems are included, but restrictive enough that the equation may be solved. Such conditions will be delimited here, thus making it possible actually to apply Booton's generalization to practical design problems.

A set of such conditions was announced earlier (4). In reference (4) there were two conditions mentioned, one of which was very reasonable, being such that it was easy to see that it would be satisfied frequently in practice. The meaning of the other condition used in (4) was not so clear. In the present paper, this second condition will be eliminated.

We shall restrict ourselves here to white noise. The method differs slightly when the noise is not white and its description under these more general circumstances will be reserved for a subsequent paper.

Since systems with nonstationary inputs are usually time-varying if they are optimum [cf. (2)], the paper begins with a very brief discussion of time-varying systems. We then turn to a consideration of the integral equation for the optimum. The solution of this equation when the noise is white is then given. Finally, we present some examples.

Without further mention, we shall consistently adhere to the following notations: Lower case letters will be used to refer to scalars, upper case to vectors. Given any vector, denoted by an upper case letter, the same letter in lower case will be used to refer to its components.

Time-Varying Systems

Since the idea of a time-varying system is perhaps not so familiar as it might be and since such systems will arise in what follows, we begin with a short presentation of the fundamental superposition principle for such systems.

Let $i(t)$ denote the input to a system. If the system is linear and time-*in*variant, the superposition principle informs us that the corresponding output may be written in the form

$$x(t) = \int_0^\infty g(\tau)i(t - \tau)d\tau = \int_{-\infty}^t g(t - \tau)i(\tau)d\tau \quad [1]$$

The function g appearing here is known as the impulse response of the system.

For systems which vary with time but which still are linear, the notion of an impulse response remains available; the superposition principle for such systems can be written as the following generalization of Equation [1]

$$x(t) = \int_{-\infty}^t g(t, \tau)i(\tau)d\tau \quad [2]$$

Thus, a linear system is always specified by its impulse response $g(t, \tau)$. If this response depends on the difference, $t - \tau$, alone, the system may be called time-invariant.

Actually, we shall concern ourselves solely with a special case of Equation [2]. It seems reasonable to assume that there always will be a certain distinguished instant when things begin to happen. Thus, no interest is attached to any time before the telephone it might be desired to design is first picked up or the missile is first fired, and so on. For this reason, we shall take all inputs to be zero before a certain time. By choosing the origin of time appropriately, this special instant of time also may be called zero. In this case Equation [2] becomes

$$x(t) = \int_0^t g(t, \tau)i(\tau)d\tau \quad [3]$$

Even though all the discussion which follows can be made to include the more general case, Equation [2], for the reason given we shall restrict ourselves entirely to Equation [3].

Integral Equation for the Optimum

In this section, we shall write down the integral equation for an optimum system. In order to do this it will be convenient first to introduce some notation.

Consider an ensemble of inputs i to a system. We shall con-

[1] Aeronautical Research Scientist, National Advisory Committee for Aeronautics, Ames Aeronautical Laboratory.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, December 27, 1956. Paper No. 57—IRD-3.

sider these inputs to be additive mixtures of messages and certain disturbances called noise. There will, of course, be an entire ensemble of messages. Denote a typical message by the symbol $m(t; P)$; the vector $P = (p_1, \ldots p_\alpha)$ indicates which message of the ensemble is being considered. Similarly, let $n(t; Q)$ denote a typical noise function. Finally, let $\mu(t; P)$ denote the desired output of the system. There will always be such an output, depending on the message; thus, if the problem is one of filtering, where it is desired to find the best approximation to the message, given the input, we should set $\mu(t; P) = m(t; P)$; if the problem is one of prediction of the message h seconds hence, $\mu(t, P) = m(t + h, P)$; and so on. We assume that some distribution of messages and noise is available; then, if $f(t; P, Q)$ is any function, by $Av\langle f(t; P, Q)\rangle$, we shall mean the average value of f with respect to the p's and q's. It should be noted that in contrast to Wiener's stationary theory where *time* averages are the only ones considered, if nonstationary problems are to be included, it will be these *ensemble* averages which will be of interest [cf. (2)].

According to what has gone before, a typical input to the system will be

$$i(t; P, Q) = m(t; P) + n(t; Q) \quad \ldots\ldots\ldots\ldots [4]$$

Now, consider a linear system with impulse response $g(t, \tau)$; that is, a device whose response to the input [4] is

$$x(t; P, Q) = \int_0^t g(t, \tau) i(\tau; P, Q) d\tau \quad \ldots\ldots\ldots [5]$$

The question we wish answered can now be formulated as follows: What is the impulse response g which will make the mean square error

$$\epsilon^2(t) \equiv Av\,\langle [\mu(t; P) - x(t; P, Q)]^2 \rangle$$

a minimum for $t \geq 0$? Note that our only concern is for $t \geq 0$, since for $t < 0$ everything (including ϵ^2) is zero.

Use of Equation [5] shows that we wish to minimize

$$\epsilon^2 = Av\left\langle \left[\mu(t; P) - \int_0^t g(t, \tau) i(\tau; P, Q) d\tau \right]^2 \right\rangle, \quad t \geq 0 \ldots [6]$$

In order to answer the question, we introduce the following correlation functions

$$\left.\begin{aligned} \varphi_{\mu\mu}(t, \tau) &= Av\,\langle \mu(t; P)\mu(\tau; P)\rangle \\ \varphi_{\mu i}(t, \tau) &= Av\,\langle \mu(t; P) i(\tau; P, Q)\rangle \\ \varphi_{ii}(t, \tau) &= Av\,\langle i(t; P, Q) i(\tau; P, Q)\rangle \end{aligned}\right\} \quad \ldots\ldots [7]$$

Assuming the averaging process and the integral in Equation [6] can be interchanged, we obtain by squaring and averaging that

$$\epsilon^2 = \varphi_{\mu\mu}(t, t) - 2\int_0^t g(t, \tau)\varphi_{\mu i}(t, \tau) d\tau + \int_0^t g(t, \tau) \int_0^t g(t, \sigma)\varphi_{ii}(\tau, \sigma) d\sigma\, d\tau, \quad t \geq 0 \ldots\ldots [8]$$

Now, by exactly the same methods as were used in references (1, 2, 3), it can be shown that the necessary and sufficient condition that g minimize ϵ^2 is that g satisfy the integral equation

$$\varphi_{\mu i}(t, \tau) = \int_0^t g(t, \sigma)\varphi_{ii}(\tau, \sigma) d\sigma, \quad \text{for } 0 \leq \tau \leq t \ldots [9]$$

Substitution of Equation [9] into [8] shows that the minimum square error is

$$\epsilon^2 = \varphi_{\mu\mu}(t, t) - \int_0^t g(t, \tau)\varphi_{\mu i}(t, \tau) d\tau \quad \ldots\ldots [10]$$

Up to this point, we have assumed nothing about the noise. As we announced in the Introduction, we shall consider here only white noise which is independent of the messages, leaving for a subsequent paper the application to more general situations. By definition of white noise [see (5)], there is a constant λ such that the autocorrelation of the noise

$$\varphi_{nn}(t, \tau) \equiv Av\,\langle n(t; Q) n(\tau; Q)\rangle = \lambda\delta(t - \tau)$$

where $\delta(t)$ denotes the Dirac δ-function (6). This means that if the noise is white and independent of the messages (so that $\varphi_{\mu n} = 0$), the Functions [7] satisfy the relations

$$\left.\begin{aligned} \varphi_{\mu i}(t, \tau) &= \varphi_{\mu m}(t, \tau) \\ \varphi_{ii}(t, \tau) &= \varphi_{mm}(t, \tau) + \lambda\delta(t - \tau) \end{aligned}\right\} \quad \ldots\ldots [11]$$

Making use of the fundamental property

$$\int_0^t g(t, \sigma)\delta(\tau - \sigma) d\sigma = g(t, \tau) \quad \text{if } 0 \leq \tau \leq t$$

of the δ-function, along with Relations [11], we see that the integral Equation [9] becomes

$$\varphi_{\mu m}(t, \tau) = \int_0^t g(t, \sigma)\varphi_{mm}(\tau, \sigma) d\sigma + \lambda g(t, \tau), \quad \text{for } 0 \leq \tau \leq t \quad \ldots\ldots [12]$$

while the Equation [10] for the minimum mean square error reduces to

$$\epsilon^2 = \varphi_{\mu\mu}(t, t) - \int_0^t g(t, \tau)\varphi_{\mu m}(t, \tau) d\tau \quad \ldots\ldots [13a]$$

We note in passing that if the problem is one of filtering, so that $\mu = m$, Equation [12] can be used to reduce Equation [13a] to

$$\epsilon^2 = \lambda g(t, t) \quad \ldots\ldots\ldots\ldots [13b]$$

Solution of the Integral Equation

In order to solve Equation [12], it is necessary to impose some conditions on the autocorrelation φ_{mm}. The first thing to notice is that this function is always symmetric—it follows immediately from its Definition [7] that

$$\varphi_{mm}(t, \tau) = \varphi_{mm}(\tau, t) \quad \ldots\ldots\ldots\ldots [14]$$

But more than this is needed. Suppose $m(t; P)$ were a continuous function of the vector P. Then, as is well known, m can be approximated as closely as desired by a polynomial in the components of P. Substitution of this polynomial approximation into the first of Equations [7] shows that in this case $\varphi_{mm}(t, \tau)$ is a sum of products of functions of t and functions of τ. Thus, with this approximation, the following statement becomes true:

There is a set of functions $a_1(t), \ldots, a_\alpha(t)$, $b_1(t), \ldots, b_\alpha(t)$, $c_1(t), \ldots, c_\alpha(t)$, such that

$$\left.\begin{aligned} \varphi_{mm}(t, \tau) &= \sum_{\rho=1}^{\alpha} a_\rho(t) b_\rho(\tau) \\ \varphi_{\mu m}(t, \tau) &= \sum_{\rho=1}^{\alpha} c_\rho(t) b_\rho(\tau) \end{aligned}\right\} \quad \text{for } t \geq \tau \ldots\ldots [15]$$

Equations [15] are very often exact. Even if they are not, the foregoing discussion shows that they are approximately true.

If the Assumption [15] is made, use of Equation [14] shows that

$$\varphi_{mm}(t, \tau) = \sum_{\rho=1}^{\alpha} a_\rho(\tau) b_\rho(t) \quad \text{for} \quad \tau > t \ldots\ldots [16]$$

Now, let A, B, and C denote the vectors $\{a_\rho\}$, $\{b_\rho\}$, and $\{c_\rho\}$, respectively. Then Equations [15] and [16] can be written in the succinct form

$$\left.\begin{aligned} \varphi_{mm}(t,\tau) &= \begin{cases} A(t)\cdot B(\tau) & \text{for } \tau \le t \\ A(\tau)\cdot B(t) & \text{for } \tau > t \end{cases} \\ \varphi_{\mu m}(t,\tau) &= C(t)\cdot B(\tau) \quad \text{for } \tau \le t \end{aligned}\right\} \quad \ldots\ldots [17]$$

where the dot denotes the ordinary scalar product. We shall make the Assumption [17] throughout this paper.

The computations which are to follow are rather complicated. Hence, it is suggested that the remainder of this section be read along with the solution of an example—those in the next section, for instance.

If Assumption [17] is substituted into [12], it can be seen that the fundamental integral equation becomes

$$C(t)\cdot B(\tau) = A(\tau)\cdot\int_0^\tau B(\sigma)g(t,\sigma)d\sigma + B(\tau)\cdot\int_\tau^t A(\sigma)g(t,\sigma)d\sigma + \lambda g(t,\tau) \quad \text{for } 0 \le \tau \le t \ldots\ldots [18]$$

Before going on with the general analysis, we stop here to consider a special case which sometimes arises and which can be solved immediately. This is the case when $B = A$. With this assumption, Equation [18] becomes

$$C(t)\cdot A(\tau) = A(\tau)\cdot\int_0^t A(\sigma)g(t,\sigma)d\sigma + \lambda g(t,\tau) \quad \text{for } 0 \le \tau \le t \ldots\ldots [19]$$

This equation is degenerate [see reference (7)] and can be solved as follows: Set

$$\int_0^t A(\sigma)g(t,\sigma)d\sigma = J(t)$$

Multiply Equation [19] by $A(\tau)$ and integrate with respect to τ from zero to t. This gives

$$\int_0^t [C(t)\cdot A(\tau)]A(\tau)d\tau = \int_0^t [J(t)\cdot A(\tau)]A(\tau)d\tau + \lambda J(t) \ldots [20]$$

Set $$\alpha_{\nu\rho}(t) = \int_0^t a_\nu(\tau)a_\rho(\tau)d\tau$$

Then, writing all vectors in terms of their components, it can be seen that Equation [20] is equivalent to

$$\lambda j_\nu(t) + \sum_\rho \alpha_{\nu\rho}(t)j_\rho(t) = \sum_\rho \alpha_{\nu\rho}(t)c_\rho(t) \ldots\ldots [21]$$

This is a system of simultaneous linear equations for the functions j_ν. Solving Equations [21] gives the vector J and hence the impulse response $g(t,\tau)$, for, from Equation [19]

$$g(t,\tau) = \frac{C(t) - J(t)}{\lambda}\cdot A(\tau), \quad t \ge \tau \ldots\ldots [22]$$

To return to the general case, set

$$v(t,\tau) \equiv A(t)\cdot B(\tau) - A(\tau)\cdot B(t) \ldots\ldots [23]$$

Then, writing $\int_\tau^t$ which occurs in Equation [18] as $\int_0^t - \int_0^\tau$, we obtain

$$\left[C(t) - \int_0^t A(\sigma)g(t,\sigma)d\sigma\right]\cdot B(\tau) = \int_0^\tau g(t,\sigma)v(\tau,\sigma)\,d\sigma + \lambda g(t,\tau) \quad \text{for } 0 \le \tau \le t \ldots\ldots [24]$$

We now attempt to find a solution of the form

$$g(t,\tau) = [G(t)\cdot\Gamma(\tau)]u(t-\tau) \ldots\ldots [25]$$

where $u(t)$ is the unit step function

$$u(t) = \begin{cases} 0, & t < 0 \\ 1, & t \ge 0 \end{cases}$$

The function $u(t-\tau)$ enters into Equation [25] since it is necessary that $g(t,\tau)$ be zero for $t < \tau$; if this were not the case, the system would be required to respond to an input before the latter occurred, which is clearly impossible for physically realizable systems.

Now, substitute Equation [25] into [24]. This gives

$$\left\{C(t) - \int_0^t A(\sigma)[G(t)\cdot\Gamma(\sigma)]d\sigma\right\}\cdot B(\tau) = G(t)\cdot\left[\int_0^\tau \Gamma(\sigma)v(\tau,\sigma)d\sigma + \lambda\Gamma(\tau)\right] \quad \text{for } 0 \le \tau \le t \ldots [26]$$

Equation [26] is most certainly satisfied if

$$\left.\begin{aligned} &(a)\quad B(\tau) = \int_0^\tau \Gamma(\sigma)v(\tau,\sigma)d\sigma + \lambda\Gamma(\tau), \quad 0 \le \tau \le t \\ &(b)\quad C(t) - \int_0^t A(\sigma)[G(t)\cdot\Gamma(\sigma)]d\sigma = G(t), \quad t \ge \tau \end{aligned}\right\} \ldots [27]$$

By deriving Equations [27], we have reduced the integral Equation [12] which depends on two independent variables to a set of equations each depending on a single such variable. It remains to solve Equations [27].

The first of Equations [27] still appears to have a very general form. However, the kernel v is the type which is a sum of products of functions of one of its variables and functions of the other variable; this can be seen from the Definition [23] of v. Since v is of this type, we shall find that it is possible to solve the equation. Indeed, define the vectors

$$\left.\begin{aligned} E(\tau) &= [a_1(\tau), \ldots, a_\alpha(\tau), b_1(\tau), \ldots, b_\alpha(\tau)] \\ F(\tau) &= [b_1(\tau), \ldots, b_\alpha(\tau), -a_1(\tau), \ldots, -a_\alpha(\tau)] \end{aligned}\right\} \ldots [28]$$

Then, according to Equation [23]

$$v(t,\tau) = E(t)\cdot F(\tau)$$

and so Equation [27a] becomes

$$b_p(\tau) = E(\tau)\cdot\int_0^\tau F(\sigma)\gamma_p(\sigma)d\sigma + \lambda\gamma_p(\tau) \quad \text{for } 0 \le \tau \le t, \quad p = 1, \ldots, \alpha \ldots\ldots [29]$$

when the vectors B and Γ are written in terms of their components. Thus

$$b_p(\tau) = \sum_{q=1}^{2\alpha} e_q(\tau)\int_0^\tau f_q(\sigma)\gamma_p(\sigma)d\sigma + \lambda\gamma_p(\tau) \quad \text{for } 0 \le \tau \le t,\ p = 1, \ldots, \alpha \ldots\ldots [30]$$

Now, it is entirely possible that the components $e_q(\tau)$ of $E(\tau)$ are not all linearly independent. In this case, certain of the terms on the right side of Equation [30] can be collected together, and this process can be continued until equations of the form

$$b_p(\tau) = \sum_{q=1}^{\beta} \epsilon_q(\tau)\int_0^\tau \varphi_q(\sigma)\gamma_p(\sigma)d\sigma + \lambda\gamma_p(\tau) \quad \text{for } 0 \le \tau \le t,\ p = 1, \ldots, \alpha \ldots\ldots [31]$$

are obtained where the functions $\epsilon_q(\tau)$ are linearly independent. Equations [31] can then be reduced immediately to a system of differential equations. In fact, differentiating Equations [31] r times gives

$$\sum_{q=1}^{\beta} \epsilon_q^{(r)}(\tau) \int_0^\tau \varphi_q(\sigma)\gamma_p(\sigma)d\sigma = b_p^{(r)}(\tau) - \lambda\gamma_p^{(r)}(\tau) - \sum_{s=1}^{r} \binom{r}{s} \sum_{q=1}^{\beta} \epsilon_q^{(r-s)}(\tau) \frac{d^{s-1}}{d\tau^{s-1}} [\varphi_q(\tau)\gamma_p(\tau)]$$
$$\text{for} \quad 0 \leq \tau \leq t, \; p = 1, \ldots, \alpha \quad \ldots\ldots [32]$$

where $\binom{r}{s}$ denotes the binomial coefficient

$$\binom{r}{s} = \frac{r!}{s!(r-s)!}$$

Equations [32] with $r = 0, 1, \ldots, \beta - 1$ represent β simultaneous equations in the β unknowns

$$\int_0^\tau \varphi_q(\sigma)\gamma_p(\sigma)d\sigma, \; q = 1, \ldots, \beta \; (p \text{ fixed})$$

Furthermore, the coefficient determinant

$$\det[\epsilon_q^{(r)}] \ldots\ldots [33]$$

is never zero, since we know the functions ϵ_q to be independent, from which it follows that their Wronskian [33] does not vanish. Consequently, Equations [32] can be solved for the integrals

$$\int_0^\tau \varphi_q(\sigma)\gamma_p(\sigma)d\sigma \ldots\ldots [34]$$

in terms of the function γ_p and their derivatives. Differentiating these solution equations once more results in a linear differential equation for $\gamma_p(\tau)$. This equation determines γ_p uniquely because (a) it is nonsingular if the vectors A and B are, say, continuous, since the coefficient of the highest derivative of γ_p is λ which is never zero, and (b) the initial conditions are determined by the fact that $\gamma_p(\tau) = 0$ for $\tau < 0$.

There is only one thing which it appears might go wrong here. With p fixed, we have determined β differential equations by differentiating the Integrals [34] for $q = 1, \ldots, \beta$. It might seem possible that these equations are different and do not possess a common solution. This cannot be the case, however, for Equation [27a] is a Volterra equation which always possesses a unique solution. Since our differential equations are implied by the integral Equation [27a], however, this means that the former must have a common solution.

Thus, we have found a set of linear differential equations for the functions γ_p. These equations may or may not be explicitly solvable in terms of the known functions of analysis. Very frequently they are, but even if they are not, the problem has been reduced from finding the solution of an integral equation to finding the solution of a linear differential equation, about which a great deal is known, even if only about approximations to solutions.

So γ_p can be found. To find the components g_p of the vector G, we return to Equation [27b]. Componentwise, this equation can be written

$$g_p(t) + \sum_{q=1}^{\alpha} g_q(t) \int_0^t a_p(\sigma)\gamma_q(\sigma)d\sigma = c_p(t), \quad p = 1, \ldots, \alpha \ldots\ldots [35]$$

Since the functions γ_q are now known, the integrals occurring here can be computed. Then, Equations [35] become a system of α simultaneous, linear, algebraic equations for the α functions $g_p(t)$.

We have now described a method for finding α functions γ_p and α functions g_p. The desired impulse response can now be found from Equation [25].

Examples

Example 1. The first example we shall consider is a much simplified[3] version of the "straightforward" problem of the gun platform considered in the Introduction. We state the problem as follows: Suppose a particle leaves the origin at some fixed time ($=0$) and moves thereafter with a constant (but unknown) speed along a given straight line. It is desired to find the best approximation to the position of the particle at any time, assuming the measurements obscured by (white) noise.

The messages (i.e., the possible particle positions) here all have the form

$$m(t; p) = pt, \quad t \geq 0$$

where p is the unknown speed of the particle. As will be seen, it will not be necessary to have even a complete statistical distribution of values of p. The mean square value of p—which we shall denote by h^2—will be all that need be known to solve the problem.

Since in this example we wish the output of our system to approximate the particle position, we set $\mu = m$ to obtain

$$\begin{aligned} \varphi_{\mu\mu}(t, \tau) &= \varphi_{mm}(t, \tau) \\ &= Av\,\langle m(t; p)m(\tau; p)\rangle \\ &= Av\,\langle pt \cdot p\tau\rangle \\ &= t\tau\, Av\,\langle p^2\rangle \end{aligned}$$

Now, whatever the distribution of particle speeds may be, the average occurring here is just the mean square speed. Hence

$$\left.\begin{aligned} \varphi_{\mu\mu}(t, \tau) &= h^2 t\tau \\ \varphi_{ii}(t, \tau) &= h^2 t\tau + \lambda\delta(t - \tau) \end{aligned}\right\} \quad \text{for} \quad t, \tau \geq 0 \ldots. [36]$$

and so the integral Equation [12] becomes

$$h^2 t\tau = h^2\tau \int_0^t \sigma g(t, \sigma)d\sigma + \lambda g(t, \tau) \quad \text{for} \quad 0 \leq \tau \leq t \ldots [37]$$

In the notation of Equation [16], we have from Equation [36] that

$$a_1(t) = ht; \quad b_1(t) = ht; \quad c_1(t) = ht \ldots\ldots [38]$$

Note that $a_1 = b_1$, and so the simpler Solution [22] may be used. In fact, call

$$j(t) = \int_0^t \sigma g(t, \sigma)d\sigma$$

Multiply Equation [37] by τ and integrate with respect to that variable from zero to t. This gives

$$h^2 t \cdot \frac{t^3}{3} = h^2 \frac{t^3}{3} j(t) + \lambda j(t)$$

that is

$$j(t) = \frac{h^2 t^4}{h^2 t^3 + 3\lambda}$$

Hence, from Equation [22]

$$\begin{aligned} g(t, \tau) &= h^2\tau \cdot \frac{t - j(t)}{\lambda} \\ &= \frac{3h^2 t\tau}{h^2 t^3 + 3\lambda}, \quad t \geq \tau \ldots\ldots [39] \end{aligned}$$

This same result can be arrived at by the general method leading

[3] These simplifications are not needed to solve the problem, actually, but it is not our purpose here to specify an optimum gun platform; we wish merely to illustrate the operation of the method described earlier.

to Equations [25], [32], and [35]. Indeed, from Equations [38] it can be seen that

$$v(t, \tau) = a_1(t)b_1(\tau) - a_1(\tau)b_1(t) = 0$$

Hence, from Equation [27a]

$$\gamma_1(\tau) = \frac{1}{\lambda} b_1(\tau)$$

$$= \frac{h}{\lambda} \tau$$

Thus

$$\int_0^t a_1(\sigma)\gamma_1(\sigma)d\sigma = \frac{h^2}{\lambda} \int_0^t \sigma^2 d\sigma$$

$$= \frac{h^2}{3\lambda} t^3$$

and so from Equation [27b]

$$g_1(t) + \frac{h^2}{3\lambda} t^3 g_1(t) = ht$$

that is

$$g_1(t) = \frac{3\lambda ht}{h^2t^3 + 3\lambda}$$

Consequently, from Equation [25]

$$g(t, \tau) = g_1(t)\gamma_1(\tau)u(t - \tau)$$

$$= \frac{3h^2t\tau}{h^2t^3 + 3\lambda} u(t - \tau)$$

which agrees with Equation [39].

Equation [13b] can be used to find the rms error at any time. In fact

$$\epsilon^2 = \lambda g(t, t)$$

$$= \frac{3\lambda h^2t^2}{h^2t^3 + 3\lambda}$$

As t grows large, we may approximate to find

$$\epsilon \approx \sqrt{\frac{3\lambda}{t}}$$

Note this implies that, by waiting long enough, this error can be made as small as desired.

This example was, of course, extremely simple, owing to the degeneracy of the Equation [37], which manifested itself in the vanishing of v. A slightly more complicated problem follows.

Example 2. For our second example, we choose one with stationary inputs. This is so that the results of the theory discussed herein and those of the Wiener theory can be compared. Now, the principal advantage of the newer theory is that it can be applied to nonstationary problems where the Wiener theory can give no answer at all. However, by allowing time-varying systems into the competition for the title of optimum, we also can utilize a more reasonable definition of error, since we allow a starting time to exist. Hence, it might be supposed that some gains are to be had by using the present theory instead of the Wiener theory even where the latter can be applied. Example 2 will illustrate this.

Specifically, we shall consider a filter designed to give the best approximation to a class of messages with autocorrelation

$$\varphi_{mm}(t, \tau) = \frac{1}{2} e^{-|t-\tau|} \quad [40]$$

(This example is considered by Wiener in reference 1, pp. 91–92.)

Since the noise is white, the integral equation satisfied by the optimum in the Wiener sense (1) is

$$\frac{1}{2} e^{-t} = \frac{1}{2} \int_0^\infty g(\tau)e^{-|t-\tau|}d\tau + \lambda g(t), \quad t \geq 0 \ldots [41]$$

(Note that this impulse response depends only on one variable, since it represents a time-invariant system.) It is easy to see that Equation [41] is satisfied by

$$g_W(t) = (\beta - 1)e^{-\beta t} \quad [42]$$

where

$$\beta^2 = \frac{1 + \lambda}{\lambda} \quad [43]$$

The impulse response [42], then, is the optimum in the Wiener sense.

The error for a time-invariant system with stationary inputs and white noise can be found from Equation [8] to be

$$\epsilon^2 = \varphi_{\mu\mu}(0) - 2 \int_0^t g(\tau)\varphi_{\mu m}(\tau)d\tau$$

$$+ \int_0^t g(\tau) \int_0^t g(\sigma)\varphi_{mm}(\sigma - \tau)d\sigma\, d\tau + \lambda \int_0^t g^2(\tau)d\tau, \quad t \geq 0$$

Hence, substituting from Equations [40] and [42] into this[4] formula, we find the error corresponding to the Wiener system to be a fairly complicated expression which as $t \to \infty$ satisfies

$$\epsilon_W^2 \to \frac{1}{\beta + 1} \quad [44]$$

where β is to be found from Equation [43].

We now compute the optimum in our sense. Comparing Equations [15] and [40], we see that we may choose

$$a_1(t) = \frac{1}{2} e^{-t}; \quad b_1(t) = e^t; \quad c_1(t) = \frac{1}{2} e^{-t}, \quad t \geq 0 \ldots [45]$$

this last because $\mu = m$. From Equation [23], then

$$v(t, \tau) = \frac{1}{2} e^{-t} \cdot e^\tau - \frac{1}{2} e^{-\tau} \cdot e^t$$

$$= -\sinh (t - \tau), \quad t, \tau \geq 0 \ldots [46]$$

Since $v(t, \tau)$ is here a function of the difference, $t - \tau$, alone, Equation [27a] can be solved by Laplace transforms; however, since the purpose of these examples is the illustration of the method, we shall not use this fact.

Using Equations [45] and [46], we see that Equation [27a] may be written

$$e^\tau = \lambda\gamma(\tau) - \int_0^\tau \gamma(\sigma) \sinh (\tau - \sigma)d\sigma, \quad 0 \leq \tau \leq t \quad [47]$$

By Equations [28], we have

$$e_1(\tau) = \frac{1}{2} e^{-\tau}r; \quad e_2(\tau) = e^\tau$$

$$f_1(\tau) = e^\tau; \qquad f_2(\tau) = -\frac{1}{2} e^{-\tau}$$

Consequently, Equation [29] becomes

[4] Note that the Wiener system here is penalized since *our* Definition [8] of the error is used; this system is not, after all, designed to minimize Equation [8]. On the other hand, if our belief is correct that Equation [8] is a more realistic expression for the error, all systems should be compared on this basis.

$$e^{\tau} = \frac{1}{2}\, e^{-\tau} \int_0^{\tau} e^{\sigma}\gamma(\sigma)d\sigma - \frac{1}{2}\, e^{\tau} \int_0^{\tau} e^{-\sigma}\gamma(\sigma)d\sigma + \lambda\gamma(\tau) \quad \text{for} \quad 0 \leq \tau \leq t \ldots\ldots [48]$$

Since, in this case, the functions e_1 and e_2 are independent, Equations [29] and [30] are identical.

Now, differentiate Equation [48]. This gives

$$e^{\tau} = -\frac{1}{2}\, e^{-\tau} \int_0^{\tau} e^{\sigma}\gamma(\sigma)d\sigma - \frac{1}{2}\, e^{\tau} \int_0^{\tau} e^{-\sigma}\gamma(\sigma)d\sigma + \lambda\dot{\gamma}(\tau) \ldots\ldots [49]$$

Equations [48] and [49] can be solved for the integrals occurring therein. In fact, we find by adding these equations, that

$$\int_0^{\tau} e^{-\sigma}\gamma(\sigma)d\sigma = \lambda e^{-\tau}(\gamma + \dot{\gamma}) - 2$$

Hence, differentiating

$$\gamma(\tau) = \lambda(\ddot{\gamma} - \gamma)$$

that is

$$\ddot{\gamma} - \beta^2\gamma = 0$$

where β^2 is given by Equation [43]. Thus, solving this equation, we find

$$\gamma(\tau) = k_1 e^{\beta\tau} + k_2 e^{-\beta\tau} \ldots\ldots [50]$$

where k_1 and k_2 are constants.

At this point, it should be noted that the same expression for γ would have been found if, instead of being added, Equations [48] and [49] had been subtracted to find

$$\int_0^{\tau} e^{\sigma}\gamma(\sigma)d\sigma$$

The values of k_1 and k_2 can be found by substitution into Equation [47]. Using Equation [50] in [47], we find

$$e^{\tau} = k_1\left[\lambda e^{\beta\tau} - \int_0^{\tau} e^{\beta\sigma}\sinh(\tau - \sigma)d\sigma\right] + k_2\left[\lambda e^{-\beta\tau} - \int_0^{\tau} e^{-\beta\sigma}\sinh(\tau - \sigma)d\sigma\right]$$

$$= k_1\left[\frac{e^{\tau}}{2(\beta - 1)} - \frac{e^{-\tau}}{2(\beta + 1)}\right] - k_2\left[\frac{e^{\tau}}{2(\beta + 1)} - \frac{e^{-\tau}}{2(\beta - 1)}\right]$$

Equating coefficients of like functions of τ on both sides of this equation, one finds

$$1 = \frac{k_1}{2(\beta - 1)} - \frac{k_2}{2(\beta + 1)}$$

$$0 = -\frac{k_1}{2(\beta + 1)} + \frac{k_2}{2(\beta - 1)}$$

That is, using Equation [43]

$$k_1 = \frac{\beta + 1}{2\beta\lambda}, \quad k_2 = \frac{\beta - 1}{2\beta\lambda}$$

so that from Equation [50]

$$\gamma(\tau) = \frac{(\beta + 1)e^{\beta\tau} + (\beta - 1)e^{-\beta\tau}}{2\beta\lambda} \ldots\ldots [51]$$

Now, by Equations [45] and [51]

$$\int_0^t a_1(\sigma)\gamma(\sigma)d\sigma = \frac{(\beta + 1)^2}{4\beta}\, e^{(\beta-1)t} - \frac{(\beta - 1)^2}{4\beta}\, e^{-(\beta+1)t} - 1$$

Hence, from Equation [35]

$$g(t) = \frac{2\beta}{(\beta + 1)^2 e^{\beta t} - (\beta - 1)^2 e^{-\beta t}}$$

Consequently, Equation [25] gives

$$g(t, \tau) = g(t)\gamma(\tau)u(t - \tau) = \frac{1}{\lambda}\, \frac{(\beta + 1)e^{\beta\tau} + (\beta - 1)e^{-\beta\tau}}{(\beta + 1)^2 e^{\beta t} - (\beta - 1)^2 e^{-\beta t}}\, u(t - \tau) \ldots\ldots [52]$$

To compute the error associated with this system, consider Equation [13b]. We have

$$\epsilon^2 = \lambda g(t, t) = \frac{(\beta + 1)e^{\beta t} + (\beta - 1)e^{-\beta t}}{(\beta + 1)^2 e^{\beta t} - (\beta - 1)^2 e^{-\beta t}}$$

so that

$$\epsilon(\infty) = 1\sqrt{(\beta + 1)}$$

which is the same as the error [44] at infinity of the Wiener system. This result is not surprising, actually, since at $t = \infty$, statistical equilibrium will have been reached regardless of the starting time. Since this equilibrium is a design condition of the Wiener system, it might be expected that no better linear system—time-varying or not—can be found. This can be looked at in another way, as follows: As t and τ move away from zero, the second terms in the numerator and denominator of $g(t, \tau)$ become negligible in comparison with the first, and so Equation [52] becomes

$$g(t, \tau) \approx (\beta - 1)e^{-\beta(t-\tau)}$$

[using the fact that $\lambda = 1/(\beta^2 - 1)$], which describes a time-invariant system—the same system, in fact, as was described by Equation [42].

For relatively small values of t, the system with impulse response [52] has, of course, a smaller associated error than does the Wiener system. Thus, we may conclude that if one is considering stationary inputs (such as might arise in meteorological prediction problems, for example), the Wiener system and the system designed by the method of this report will give the same results unless one is concerned with short runs, in which case the present method is somewhat superior. If the inputs are nonstationary, the Wiener method can, of course, not be used at all, while the present method remains available.

BIBLIOGRAPHY

1 "Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications," by Norbert Wiener, John Wiley & Sons, Inc., New York, N. Y., 1949.

2 "An Optimization Theory for Time-Varying Linear Systems with Nonstationary Statistical Inputs," by R. C. Booton, Jr., MIT Dynamic Analysis and Control Laboratory, Cambridge, Mass., July, 1951, Meteor Report 72.

3 "A Heuristic Exposition of Wiener's Mathematical Theory of Prediction and Filtering," by Norman Levinson, *Journal of Mathematics and Physics*, vol. 26, 1947, pp. 110–119.

4 "On a Method for Optimization of Time-Varying Linear Systems with Nonstationary Inputs," by Marvin Shinbrot, NACA TN 3791, 1956.

5 "Theory of Servomechanisms," by H. M. James, N. B. Nichols, and R. S. Phillips, McGraw-Hill Book Company, Inc., New York, N. Y., 1947.

6 "The Principles of Quantum Mechanics," by P. A. M. Dirac, Clarendon Press, Oxford, England, third edition, 1947.

7 "Methoden der Mathematischen Physik," by R. Courant and D. Hilbert, Julius Springer, Berlin, Germany, 1931.

Design of Multivariable Optimum Filters

By J. H. WESTCOTT,[1] LONDON, ENGLAND

The paper considers in detail a case of multivariable optimum filter design which is of engineering interest. This is the problem of extracting the best resemblance, in a minimum mean-square-error sense, of a message available in differently corrupted forms from a number of sources, given the statistical characteristics of message and disturbances. The solution is shown to involve an essential difference from the familiar case for a single source. Other multivariable optimum studies are not in principle different from the one considered here and consequently require the same type of analysis. A numerical example of the design of the optimum combination of filters for deriving a message from two noisy sources is given.

Introduction

THE classical single-channel filter used in communication systems has an extensive literature of its own. It is concerned with the technique of subdividing the channel band width into sharply defined packages each of which is required to have minimum distortion in the pass band, to have minimum overlap as between packages, and to be suitably excluding to all frequencies occupied by other pass bands.

Recently the term "filter" has been used in a broader sense to apply to cases in which the filter characteristic is assessed from statistical properties of the signals. A filter in this sense may be a complicated piece of equipment with, for example, a memory store containing a given set of possible transmitted messages which are used to enable a particular message from the set to be recognized with minimum error on the average, although the message has been corrupted by random disturbance in the transmission process. This type of filter has been discussed by Fano.[2]

A further example of this use of the word in a broader sense is the filter discussed by Wiener,[3] in which only the statistical properties of the generators of the message and corrupting disturbance are known. The filter is then required to recover the original message with minimum mean-square error. Wiener[4] also discusses an extension to this case which is analogous to Fano's filter in that it is required to recognize specific messages from sets of messages, but the generation statistics of the messages only are known. In this respect it is a more general case than the one considered by Fano. Unfortunately, the discussion is difficult to follow, and is not made any easier by the presence of misprints in the text, some of which are very misleading. The essential idea is the use of the method of undefined coefficients which also is used in the present paper for a different application. This application is felt to be of more direct interest to control engineers and is that of finding the best set of filters for extracting a message which is available from several sources, each source involving an independent corrupting disturbance. The work is most straightforward for the simple case of two channels in which both disturbance signals are independent of the message. An example is given in the paper for this case. In principle, there is no difference for disturbances having cross correlation with the message or for cases involving more than two channels. A complete analysis is given in an Appendix. However, even the simple case of two channels is different in principle from the now familiar case of a single channel since a direct explicit solution is no longer possible.

Optimum Extraction of a Message From Two Noisy Channels

The simplest case of the problem, namely, for two channels, is illustrated in Fig. 1. The message $m(t)$ is available in both channels but is corrupted by disturbances $n(t)$ in the one case and $\nu(t)$ in the other. These signals are known only in statistical terms; they are assumed to possess auto-correlation and cross-correlation functions that can be assessed by measurement. The criterion for the optimum result is minimization of mean-square error between the summed output from the two channels and the required message. The method of determining the filters is to express the mean-square error in terms of the two filter transfer functions $H_1(\omega)$ and $H_2(\omega)$ and to minimize with respect to the form of these filter frequency characteristics. A complete analysis for the two-channel case in which message and disturbances are all correlated is given in the Appendix. The essential steps in this analysis are more easily followed for the restricted case in which all the signals are independent and so have no cross correlations. Following the procedure used in the Appendix of expressing the mean-square error in terms of time functions, then changing the order of integration in order to substitute for correlation functions, and finally taking frequency transforms, the mean-square error $\overline{e^2}$ is given in terms of the filter transfer functions $H_1(\omega)$ and $H_2(\omega)$ by the following expression

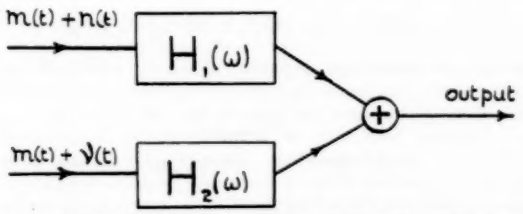


Fig. 1 Extraction of a Message From Two Noisy Channels

$$\overline{e^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \{[\Phi_m(\omega) + \Phi_n(\omega)]|H_1(\omega)|^2 + [\Phi_m(\omega) + \Phi_\nu(\omega)]|H_2(\omega)|^2 + \Phi_m(\omega) + \Phi_m(\omega)[H_1(\omega)\bar{H}_2(\omega) + \bar{H}_1(\omega)H_2(\omega)] - \Phi_m(\omega)e^{j\omega a}[\bar{H}_1(\omega) + \bar{H}_2(\omega)] - \Phi_m(\omega)e^{-j\omega a}[H_1(\omega) + H_2(\omega)]\}d\omega \quad \ldots\ldots [1][5]$$

---

[1] Imperial College.

[2] "Communication in the Presence of Additive Gaussian Noise," by R. M. Fano; "Communication Theory," edited by Dr. Willis Jackson, Butterworths Scientific Publication, London, England, 1952.

[3] "The Extrapolation, Interpolation, and Smoothing of Stationary Time Series," by N. Wiener, John Wiley & Sons, Inc., New York, N. Y., 1949.

[4] Loc. cit., chapter IV, p. 104.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-11.

[5] Since there is no cross correlation involved only one suffix to the Φ's is required. The bars over the H's signify "conjugate of."

The time-delay factor $e^{-j\omega\alpha}$ is introduced for the case where a time lapse before the arrival of the message is allowable, if by this concession a better filter can be obtained. The next step in the analysis involves the use of the calculus of variation and enables the condition for $\overline{e^2}$ a minimum to be obtained with respect to variation in the form of the network functions $H_1(\omega)$ and $H_2(\omega)$. The condition reduces to the following pair of equations which must be satisfied simultaneously

$$[\Phi_m(\omega) + \Phi_n(\omega)]H_1(\omega) + \Phi_m(\omega)H_2(\omega) - \Phi_m(\omega)e^{j\omega\alpha} = Q_1(\omega) \quad [2]$$

$$\Phi_m(\omega)H_1(\omega) + [\Phi_m(\omega) + \Phi_v(\omega)]H_2(\omega) - \Phi_m(\omega)e^{j\omega\alpha} = Q_2(\omega) \quad [3]$$

in which $Q_1(\omega)$ and $Q_2(\omega)$ are functions having poles in the lower half plane only, ω being regarded as a complex variable. At this stage all the Φ's are known, but $Q_1(\omega)$ and $Q_2(\omega)$ are not known, and in fact, it is not necessary to know them in detail as will be seen; all that it is necessary to know about them is that their pole positions have the special property of lying in the lower half plane only. It is the need to deal with a pair of simultaneous relationships between $H_1(\omega)$ and $H_2(\omega)$ as the condition for minimizing $\overline{e^2}$ which makes this case essentially different from the single-channel case. In order to continue using the familiar methods for solving for simultaneous algebraic equations, it is necessary to make these conditions into a pair of equations by the introduction of functions $Q_1(\omega)$ and $Q_2(\omega)$ whose special properties are for the moment ignored. Using Cramer's rule then gives

$$\Phi(\omega)H_1(\omega) = \Phi_m(\omega)\Phi_v(\omega)e^{j\omega\alpha} + Q_1(\omega)[\Phi_m(\omega) + \Phi_v(\omega)] - Q_2(\omega)\Phi_m(\omega) \quad [4]$$

where

$$\Phi(\omega) = [\Phi_m(\omega) + \Phi_n(\omega)][\Phi_m(\omega) + \Phi_v(\omega)] - \Phi_m^2(\omega)$$

Let $\Phi(\omega)$ be divided into the product of two factors $\Phi^+(\omega)\Phi^-(\omega)$ such that $\Phi^+(\omega)$ has zeros and poles in the upper half plane of ω only, and $\Phi^-(\omega)$ has zeros and poles in the lower half plane of ω only. Dividing both sides of Equation [4] by $\Phi^-(\omega)$ gives $H_1(\omega)\Phi^+(\omega)$ on the left-hand side. Since $H_1(\omega)$ is required to be a physically realizable filter, it is required to have poles in the upper half plane only as indeed has $\Phi^+(\omega)$. Use is made of this contrived circumstance since it is then only necessary to take that part of the right-hand side of Equation [4] arising from the residues at upper half-plane poles in order to have a satisfactory equation: Thus

$$H_1(\omega)\Phi^+(\omega) = \left[\frac{\Phi_m(\omega)\Phi_v(\omega)e^{j\omega\alpha}}{\Phi^-(\omega)}\right]_+ + \sum_{r=1}^{j}\frac{a_r}{\omega - \omega_r} + \sum_{s=1}^{k}\frac{b_s}{\omega - \omega_s} \quad [5]^6$$

where ω_r are upper half-plane poles of

$$\left[\frac{Q_1(\omega)(\Phi_m(\omega) + \Phi_v(\omega))}{\Phi^-(\omega)}\right]$$

and a_r are the residues at ω_r of this function and ω_s are upper half-plane poles of

$$\left[\frac{-Q_2(\omega)\Phi_m(\omega)}{\Phi^-(\omega)}\right]$$

and b_s are the residues at ω_s of this function. Since $Q_1(\omega)$ and $Q_2(\omega)$ are not known, the coefficients a_r and b_s are undefined coefficients, in terms of which we can now express $H_1(\omega)$

$$H_1(\omega) = \frac{1}{\Phi^+(\omega)}\left[\frac{\Phi_m(\omega)\Phi_v(\omega)e^{j\omega\alpha}}{\Phi^-(\omega)}\right]_+ + \frac{1}{\Phi^+(\omega)}\sum_{r=1}^{j}\frac{a_r}{\omega - \omega_r} + \frac{1}{\Phi^+(\omega)}\sum_{s=1}^{k}\frac{b_s}{\omega - \omega_s} \quad [6]$$

Similarly, by substituting for $H_1(\omega)$ in Equation [3] gives $H_2(\omega)$ also in terms of coefficients a_r and b_s. But by substituting back into Equation [2] for both $H_1(\omega)$ and $H_2(\omega)$ an equation is obtained whose partial fraction expansion must be such that the residues at upper half-plane poles must be zero and hence sufficient relationships are given for the coefficients a_r and b_s to be determined uniquely. In this manner the undefined coefficients which have been carried through the analysis finally are resolved, yielding the complete solution for the network characteristics $H_1(\omega)$ and $H_2(\omega)$. Without the hypothesis of unresolved coefficients arising from the introduction of functions $Q_1(\omega)$ and $Q_2(\omega)$, use could not be made of Cramer's rule to solve the simultaneous pair of relationships between $H_1(\omega)$ and $H_2(\omega)$; their use is fundamental to the success of the method and it is unlikely that a solution in a closed form is possible.

An Example

As a simple example of the procedure, consider the case illustrated in Fig. 1 for two channels in which all signals are independent, and the disturbing noise in each case has a spectrum uniform with frequency, that is to say, the characteristics of white noise. Let

$$\Phi_m(\omega) = \frac{1}{\omega^2 + 1}; \quad \Phi_n(\omega) = \frac{1}{2}; \quad \Phi_v(\omega) = \frac{1}{4}; \quad \alpha = 0$$

hence

$$\Phi(\omega) = [\Phi_m(\omega) + \Phi_n(\omega)][\Phi_m(\omega) + \Phi_v(\omega)] - \Phi_m^2(\omega)$$

$$= \left[\frac{\omega^2 + 3}{2(\omega^2 + 1)}\right]\left[\frac{\omega^2 + 5}{4(\omega^2 + 1)}\right] - \frac{1}{(\omega^2 + 1)^2}$$

$$= \left[\frac{j\omega + \sqrt{7}}{\sqrt{8}(j\omega + 1)}\right]\left[\frac{-j\omega + \sqrt{7}}{\sqrt{8}(-j\omega + 1)}\right] = \Phi^+(\omega)\Phi^-(\omega)$$

Determination of $H_1(\omega)$. The only upper half-plane pole in both

$$\frac{Q_1(\omega)[\Phi_m(\omega) + \Phi_v(\omega)]}{\Phi^-(\omega)} \quad \text{and} \quad \frac{Q_2(\omega)\Phi_m(\omega)}{\Phi^-(\omega)}$$

is at $\omega = j$; thus a single coefficient for both the a and b is sufficient; thus

$$H_1(\omega) = \frac{1}{\Phi^+(\omega)}\left[\frac{\Phi_m(\omega)\Phi_v(\omega)e^{j\omega\alpha}}{\Phi^-(\omega)}\right]_+ + \frac{1}{\Phi^+(\omega)}\frac{a_1}{(j\omega + 1)} \quad [7]$$

Now

$$\left[\frac{\Phi_m(\omega)\Phi_v(\omega)}{\Phi^-(\omega)}\right]_+ = \frac{1}{\sqrt{2}\,(1 + \sqrt{7})}\frac{1}{(j\omega + 1)}$$

hence

$$H_1(\omega) = \left(\frac{2}{1 + \sqrt{7}} + \sqrt{8}a_1\right)\frac{1}{(j\omega + \sqrt{7})}$$

Determination of $H_2(\omega)$. From Equation [3] using residues at upper half-plane poles only

$$H_2(\omega) = \frac{1}{[\Phi_m(\omega) + \Phi_v(\omega)]^+}\left[\frac{\Phi_m(\omega)e^{j\omega\alpha} - \Phi_m(\omega)H_1(\omega)}{[\Phi_m(\omega) + \Phi_v(\omega)]^-}\right]_+ \quad [8]$$

[6] The symbol $[\ ]_+$ indicates separation of fractions having poles in the upper half plane after making a partial fraction expansion.

$$= \frac{2(j\omega + 1)}{(j\omega + \sqrt{5})} \left[\frac{\frac{1}{\omega^2 + 1}[1 - H_1(\omega)]}{\frac{(-j\omega + \sqrt{5})}{2(-j\omega + 1)}} \right]_+$$

For the bracketed term

$$\left[\frac{1 - H_1(\omega)}{(j\omega + 1)(-j\omega + \sqrt{5})} \right]_+ = \frac{r_1}{j\omega + 1} + \frac{r_2}{j\omega + \sqrt{7}}$$

where

$$r_1 = \frac{4}{3(1 + \sqrt{5})} + \frac{4\sqrt{(2)}a_1}{(1 + \sqrt{5})(1 - \sqrt{7})}$$

$$r_2 = \frac{2}{3(\sqrt{7} + \sqrt{5})} + \frac{4\sqrt{(2)}a_1}{(\sqrt{7} - 1)(\sqrt{7} + \sqrt{5})}$$

hence

$$H_2(\omega) = \frac{4}{(1 + \sqrt{7})} \frac{1}{(j\omega + \sqrt{7})} + \frac{8\sqrt{(2)}a_1 \left(j\omega + \frac{6}{\sqrt{7} - 1} + \sqrt{5} \right)}{(j\omega + \sqrt{5})(j\omega + \sqrt{7})}$$

Having obtained expressions for both $H_1(\omega)$ and $H_2(\omega)$ in terms of the coefficient a_1 it is now necessary to substitute both into Equation [2] and to make a partial fraction development for upper half-plane poles. Substitution for $H_1(\omega)$ and $H_2(\omega)$ gives for Equation [2]

$$\frac{\omega^2 + 3}{2(\omega^2 + 1)} \frac{2}{(1 + \sqrt{7})(j\omega + \sqrt{7})} + \frac{1}{(\omega^2 + 1)} \left[\frac{4}{(1 + \sqrt{7})(j\omega + \sqrt{7})} + \frac{8\sqrt{(2)}a_1 \left(j\omega + \frac{6}{\sqrt{7} - 1} + \sqrt{5} \right)}{(j\omega + \sqrt{5})(j\omega + \sqrt{7})} \right] - \frac{1}{\omega^2 + 1} = Q_1(\omega)$$

This expression has upper half-plane poles at $\omega = j\sqrt{5}, j\sqrt{7}$, and j and the sum of the residues at each of these poles must be zero. In particular, there will be only one factor in the partial fraction expansion having a pole at $\omega = j\sqrt{5}$, whose residue will be proportional to a_1; consequently a_1 must be zero. So finally we have

$$\left. \begin{aligned} H_1(\omega) &= \frac{2}{1 + \sqrt{7}} \frac{1}{(j\omega + \sqrt{7})} \\ H_2(\omega) &= \frac{4}{1 + \sqrt{7}} \frac{1}{(j\omega + \sqrt{7})} \end{aligned} \right\} \quad \ldots\ldots\ldots [9]$$

By substituting these solutions back into Equations [2] and [3] it is easily seen that the functions $Q_1(\omega)$ and $Q_2(\omega)$ are the same. Consequently for this case a much easier solution is given[7] by subtracting Equation [3] from Equation [2] which gives $H_1(\omega)$ in terms of $H_2(\omega)$ directly; thus

$$\Phi_n(\omega)H_1(\omega) - \Phi_s(\omega)H_2(\omega) = 0$$

or

$$H_2(\omega) = \frac{\Phi_n(\omega)}{\Phi_s(\omega)} H_1(\omega) = 2H_1(\omega)$$

[7] The author is indebted to Dr. R. N. A. Plimmer for bringing this to his attention.

substituting this back in Equation [2] gives

$$H_1(\omega) \frac{\omega^2 + 7}{2(\omega^2 + 1)} - \frac{1}{(\omega^2 + 1)} = Q_1(\omega)$$

Dividing both sides by

$$\left(\frac{-j\omega + \sqrt{7}}{-j\omega + 1} \right)$$

gives

$$\frac{H_1(\omega)}{2} \frac{j\omega + \sqrt{7}}{j\omega + 1} - \frac{1}{(-j\omega + \sqrt{7})(j\omega + 1)} = Q_1(\omega) \frac{(-j\omega + 1)}{(-j\omega + \sqrt{7})}$$

where the right-hand side has lower half-plane poles only. Thus for $H_1(\omega)$ realizable

$$H_1(\omega) = \frac{2(j\omega + 1)}{(j\omega + \sqrt{7})} \left[\frac{1}{(-j\omega + \sqrt{7})(j\omega + 1)} \right]_+ = \frac{2}{1 + \sqrt{7}} \frac{1}{(j\omega + \sqrt{7})}$$

and

$$H_2(\omega) = 2H_1(\omega)$$

The attenuation-log frequency characteristics of the two filters are shown in Fig. 3.

Comparison With Filters for Individual Isolated Channels

It is interesting to compare these filters with those that would be obtained as optimum for the two channels considered individually as illustrated in Fig. 2. For the first channel the condition for minimum mean-square error is that

$$[\Phi_m(\omega) + \Phi_n(\omega)]H_1(\omega) - \Phi_m(\omega)e^{j\omega\alpha}$$

shall have no upper half-plane poles, which gives

$$H_1(\omega) = \frac{1}{[\Phi_m(\omega) + \Phi_n(\omega)]^+} \left[\frac{\Phi_m(\omega)e^{j\omega\alpha}}{[\Phi_m(\omega) + \Phi_n(\omega)]^-} \right]_+ \quad \ldots [10]$$

By analogy for the second channel

$$H_2(\omega) = \frac{1}{[\Phi_m(\omega) + \Phi_\nu(\omega)]^+} \left[\frac{\Phi_m(\omega)e^{j\omega\alpha}}{[\Phi_m(\omega) + \Phi_\nu(\omega)]^-} \right]_+ \quad \ldots [11]$$

Substituting figures gives

$$H_1(\omega) = \frac{2}{1 + \sqrt{3}} \frac{1}{(j\omega + \sqrt{3})}; \quad H_2(\omega) = \frac{4}{1 + \sqrt{5}} \frac{1}{(j\omega + \sqrt{5})}$$

These individual filter characteristics are shown also in Fig. 3 for comparison with those characteristics which give the best filtered combination of the two channels.

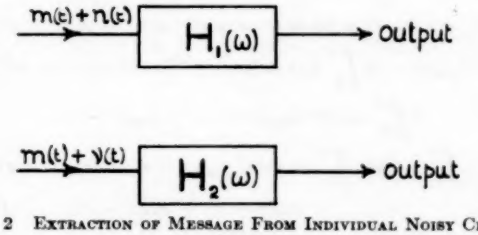


Fig. 2 Extraction of Message From Individual Noisy Channels

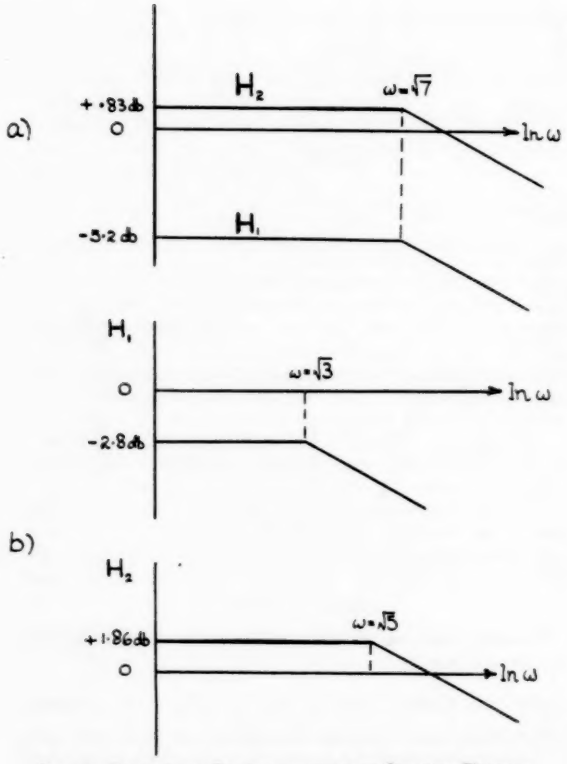


FIG. 3 FREQUENCY CHARACTERISTICS OF OPTIMUM FILTERS (*a*—For combined channels; *b*—for separate channels.)

Appendix

A message derived from two correlated noisy sources in which the noise is correlated with the message in each channel.—The mean-square error $\overline{e^2}$ may be written in terms of the message $m(t)$ and the disturbances in the two channels $n(t)$ and $v(t)$ as in Fig. 1

$$\overline{e^2} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left\{ \int_0^\infty [m(t-\tau) + n(t-\tau)] h_1(\tau) d\tau + \int_0^\infty [m(t-\tau) + v(t-\tau)] h_2(\tau) d\tau - m(t-\alpha) \right\}^2 dt \quad [12]$$

Multiplying out and reversing orders of integration gives $\overline{e^2}$ in terms of correlation functions and the weighting functions of the networks $h_1(t)$ and $h_2(t)$

$$\begin{aligned}\overline{e^2} = {} & \int_0^\infty h_1(\tau) d\tau \int_0^\infty h_1(\sigma) d\sigma [\varphi_{mm}(\tau - \sigma) + \varphi_{mn}(\tau - \sigma) \\ & + \varphi_{nm}(\tau - \sigma) + \varphi_{nn}(\tau - \sigma)] + \int_0^\infty h_2(\tau) d\tau \int_0^\infty h_2(\sigma) d\sigma \\ & [\varphi_{mm}(\tau - \sigma) + \varphi_{mv}(\tau - \sigma) + \varphi_{vm}(\tau - \sigma) + \varphi_{vv}(\tau - \sigma)] \\ & + \varphi_{mm}(0) - 2 \int_0^\infty h_1(\tau) d\tau [\varphi_{mm}(\tau - \alpha) + \varphi_{mn}(\tau - \alpha)] \\ & - 2 \int_0^\infty h_2(\tau) d\tau [\varphi_{mm}(\tau - \alpha) + \varphi_{mv}(\tau - \alpha)] \\ & + 2 \int_0^\infty h_1(\tau) d\tau \int_0^\infty h_2(\sigma) d\sigma [\varphi_{mm}(\tau - \sigma) + \varphi_{mn}(\tau - \sigma) \\ & + \varphi_{vm}(\tau - \sigma) + \varphi_{vn}(\tau - \sigma)] \quad [13]\end{aligned}$$

Taking the Fourier transform of this expression gives

$$\begin{aligned}\overline{e^2} = {} & \frac{1}{2\pi} \int_{-\infty}^{\infty} [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{nm}(\omega) + \Phi_{nn}(\omega)] |H_1(\omega)|^2 \\ & + [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{vm}(\omega) + \Phi_{vv}(\omega)] |H_2(\omega)|^2 \\ & + \Phi_{mm}(\omega) - [\Phi_{mm}(\omega) + \Phi_{mn}(\omega)] e^{-j\omega\alpha} H_1(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{nm}(\omega)] e^{j\omega\alpha} \overline{H_1(\omega)} - [\Phi_{mm}(\omega) + \Phi_{mv}(\omega)] e^{-j\omega\alpha} H_2(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{vm}(\omega)] e^{j\omega\alpha} \overline{H_2(\omega)} \\ & + [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{vm}(\omega) + \Phi_{vn}(\omega)] H_1(\omega) \overline{H_2(\omega)} \\ & + [\Phi_{mm}(\omega) + \Phi_{nm}(\omega) + \Phi_{mv}(\omega) + \Phi_{nv}(\omega)] \overline{H_1(\omega)} H_2(\omega) \quad [14]\end{aligned}$$

Let $\eta_1(\omega)$ be the variation in $H_1(\omega)$ and $\eta_2(\omega)$ be the variation in $H_2(\omega)$ then

$$\begin{aligned}\left.\frac{\partial \overline{e^2}}{\partial \epsilon}\right|_{\epsilon \to 0} = {} & \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \eta_1(\omega) \{ [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{nm}(\omega) \\ & + \Phi_{nn}(\omega)] \overline{H_1(\omega)} + [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{vm}(\omega) + \Phi_{vn}(\omega)] \overline{H_2(\omega)} \\ & - [\Phi_{mm}(\omega) + \Phi_{mn}(\omega)] e^{-j\omega\alpha} \} + \eta_2(\omega) \{ [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) \\ & + \Phi_{vm}(\omega) + \Phi_{vv}(\omega)] \overline{H_2(\omega)} + [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{nm}(\omega) \\ & + \Phi_{nv}(\omega)] \overline{H_1(\omega)} - [\Phi_{mm}(\omega) + \Phi_{mv}(\omega)] e^{-j\omega\alpha} \} \\ & + \overline{\eta_1(\omega)} \{ [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{nm}(\omega) + \Phi_{nn}(\omega)] H_1(\omega) \\ & + [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{nm}(\omega) + \Phi_{nv}(\omega)] H_2(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{nm}(\omega)] e^{j\omega\alpha} \} \\ & + \overline{\eta_2(\omega)} \{ [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{vm}(\omega) + \Phi_{vv}(\omega)] H_2(\omega) \\ & + [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{vm}(\omega) + \Phi_{vn}(\omega)] H_1(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{vm}(\omega)] e^{j\omega\alpha} \} \quad [15]\end{aligned}$$

This expression will be zero provided the terms depending on $\overline{\eta_1(\omega)}$ and $\overline{\eta_2(\omega)}$ have no upper half-plane poles, that is when

$$\begin{aligned}& [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{nm}(\omega) + \Phi_{nn}(\omega)] H_1(\omega) \\ & + [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{nm}(\omega) + \Phi_{nv}(\omega)] H_2(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{nm}(\omega)] e^{j\omega\alpha} = Q_1(\omega) \quad [16]\end{aligned}$$

$$\begin{aligned}& [\Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{vm}(\omega) + \Phi_{vv}(\omega)] H_2(\omega) \\ & + [\Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{vm}(\omega) + \Phi_{vn}(\omega)] H_1(\omega) \\ & - [\Phi_{mm}(\omega) + \Phi_{vm}(\omega)] e^{j\omega\alpha} = Q_2(\omega) \quad [17]\end{aligned}$$

where $Q_1(\omega)$ and $Q_2(\omega)$ have lower half-plane poles only. This may be written in the form

$$\Phi_{11}(\omega) H_1(\omega) + \Phi_{12}(\omega) H_2(\omega) = \Phi_{01}(\omega) e^{j\omega\alpha} + Q_1(\omega) \quad [18]$$

$$\Phi_{21}(\omega) H_1(\omega) + \Phi_{22}(\omega) H_2(\omega) = \Phi_{02}(\omega) e^{j\omega\alpha} + Q_2(\omega) \quad [19]$$

where

$$\begin{aligned}\Phi_{11}(\omega) &= \Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{nm}(\omega) + \Phi_{nn}(\omega) \\ \Phi_{12}(\omega) &= \Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{nm}(\omega) + \Phi_{nv}(\omega) \\ \Phi_{21}(\omega) &= \Phi_{mm}(\omega) + \Phi_{mn}(\omega) + \Phi_{vm}(\omega) + \Phi_{vn}(\omega) \\ \Phi_{22}(\omega) &= \Phi_{mm}(\omega) + \Phi_{mv}(\omega) + \Phi_{vm}(\omega) + \Phi_{vv}(\omega) \\ \Phi_{01}(\omega) &= \Phi_{mm}(\omega) + \Phi_{nm}(\omega) \\ \Phi_{02}(\omega) &= \Phi_{mm}(\omega) + \Phi_{vm}(\omega)\end{aligned}$$

Solving the simultaneous equations for $H_1(\omega)$ using Cramer's rule gives

$$H_1(\omega) = \frac{\begin{vmatrix} \Phi_{01}(\omega) e^{j\omega\alpha} + Q_1(\omega), \Phi_{12}(\omega) \\ \Phi_{02}(\omega) e^{j\omega\alpha} + Q_2(\omega), \Phi_{22}(\omega) \end{vmatrix}}{\begin{vmatrix} \Phi_{11}(\omega), \Phi_{12}(\omega) \\ \Phi_{21}(\omega), \Phi_{22}(\omega) \end{vmatrix}}$$

$$= \frac{[\Phi_{01}(\omega)e^{j\omega\alpha} + Q_1(\omega)]\Phi_{22}(\omega) - [\Phi_{02}(\omega)e^{j\omega\alpha} + Q_2(\omega)]\Phi_{12}(\omega)}{\Phi_{11}(\omega)\Phi_{22}(\omega) - \Phi_{12}(\omega)\Phi_{21}(\omega)} \quad [20]$$

Let

$$\Phi_{11}(\omega)\Phi_{22}(\omega) - \Phi_{12}(\omega)\Phi_{21}(\omega) = \Phi(\omega) = \Phi^+(\omega)\Phi^-(\omega) \quad [21]$$

where $\Phi^+(\omega)$ has poles and zeros in the upper half plane only and $\Phi^-(\omega)$ has poles and zeros in the lower half plane only. Then multiplying both sides of Equation [20] by $\Phi^+(\omega)$ will give an equation in which the left-hand side is required to have upper half-plane poles only; consequently, only residues at upper half-plane poles of the right-hand side are relevant to satisfying the minimum condition for ϵ^2; thus

$$H_1(\omega)\Phi^+(\omega) = \left[\frac{[\Phi_{01}(\omega)\Phi_{22}(\omega) - \Phi_{02}(\omega)\Phi_{12}(\omega)]e^{j\omega\alpha}}{\Phi^-(\omega)}\right]_+ + \sum_{r=1}^{j} \frac{a_r}{\omega - \omega_r} + \sum_{s=1}^{k} \frac{b_s}{\omega - \omega_s} \quad [22]$$

where only the residues and poles in the upper half plane of the first term on the right-hand side of Equation [22] are taken (signified by the lower limit cross outside the bracket) and ω_r are upper half-plane poles of

$$\left[\frac{Q_1(\omega)\Phi_{22}(\omega)}{\Phi^-(\omega)}\right]$$

a_r are residues at ω_r of this function; ω_s are upper half-plane poles of

$$\left[\frac{-Q_2(\omega)\Phi_{12}(\omega)}{\Phi^-(\omega)}\right]$$

and b_s are residues at ω_s of this function.

Since $Q_1(\omega)$ and $Q_2(\omega)$ are not known a_r and b_s are undefined coefficients at this stage. Thus

$$H_1(\omega) = \frac{1}{\Phi^+(\omega)}\left[\frac{[\Phi_{01}(\omega)\Phi_{22}(\omega) - \Phi_{02}(\omega)\Phi_{12}(\omega)]e^{j\omega\alpha}}{\Phi^-(\omega)}\right]_+ + \frac{1}{\Phi^+(\omega)}\sum_{r=1}^{j} \frac{a_r}{\omega - \omega_r} + \frac{1}{\Phi^+(\omega)}\sum_{s=1}^{k} \frac{b_s}{\omega - \omega_s} \quad [23]$$

Also from Equation [19] using a similar line of argument

$$H_2(\omega) = \frac{1}{\Phi_{22}{}^+(\omega)}\left[\frac{\Phi_{02}(\omega)e^{j\omega\alpha} - H_1(\omega)\Phi_{21}(\omega)}{\Phi_{22}{}^-(\omega)}\right]_+ \quad [24]$$

Substituting for $H_1(\omega)$ and $H_2(\omega)$ in Equation [18] now gives rise to an equation whose partial fraction development must be such that the residues at upper half-plane poles are zero; consequently, sufficient relationships are provided to solve for the coefficients a_r and b_s.

Discussion

RUFUS OLDENBURGER.[8] The example in the paper, showing that the "optimum" filters for two channels considered individually are different from the optimum filters when the outputs of these filters are added, is most enlightening. Nevertheless, as can be seen from Fig. 3 the difference is small; in fact, for engineering purposes the break points occur at about the same frequency. One is naturally led to ask whether or not this difference is always small, and in particular, whether corresponding break points in simple cases can be separated by an order of magnitude.

The factoring of $\Phi(\omega)$ into $\Phi^+(\omega)\Phi^-(\omega)$ may in practice be a most difficult one to do explicitly. Does the author recommend substituting simple approximations for $\Phi(\omega)$ which can be so factored?

OTTO J. M. SMITH.[9] The symbol $[\ ]_+$ used in Equation [5] and defined in footnote 6 in the paper is the realizability operator equal to the Laplace transform of the inverse Fourier transform. The characteristics of this $\mathfrak{LF}^{-1}$ operator are derived and discussed in "Separating Information from Noise," by Otto J. M. Smith. Transactions of the Professional Group on Circuit Theory, Institute of Radio Engineers, PGCT-1, December, 1952.

AUTHOR'S CLOSURE

Professor Oldenburger raises a difficulty of long standing in this type of work, namely, how may one factorize $\Phi(\omega)$ into its component product terms $\Phi^+(\omega)$ and $\Phi^-(\omega)$ when $\Phi(\omega)$ is known only for real values of ω? This is the case when $\Phi(\omega)$ has been obtained either by direct measurement or by Fourier transformation from its correlation function. A number of techniques for doing this have been discussed in the literature. The simplest are based upon curve-matching techniques using Bode plots and a set of templates for different values of relative damping ratio ζ. Simple solutions obtained in this manner can be readily improved upon using Linvill's method.[10] A general discussion of this type of procedure is given in the book by Truxal.[11] The most refined method known to the author is due to Kautz[12] and consists in the use of a generalized orthogonal set of functions. In fact Kautz offers a wide range of possibilities in choosing an orthogonal set, so that an element of skill is still required in order to get the simplest good approximation. Other methods of approximation depend on comparing coefficients or choosing coefficients so as to minimize a measure. A simple method of this last sort is discussed by Schumacher[13] although here it is difficult to see what connection there is between the measure minimized and practical performance. It is unlikely that the last word has been said on this problem, but sufficient work has already been done for a small catalog of methods to exist from which one can be selected to suit the circumstances of the problem, and the accuracy required.

[8] Professor of Mechanical and Electrical Engineering, School of Mechanical Engineering, Purdue University, Lafayette, Ind. Mem. ASME.

[9] Associate Professor of Electrical Engineering, University of California, Berkeley, Calif.

[10] "The Selection of Network Functions to Approximate Prescribed Frequency Characteristics," by J. C. Linvill, M.I.T. Research Laboratory of Electronics Technical Report No. 145, March, 1950.

[11] "Automatic Feedback Control System Synthesis," by J. G. Truxal, McGraw-Hill Publishing Co., New York, N. Y., 1955.

[12] "Network Synthesis for Specified Transient Response," by W. H. Kautz, M.I.T. Research Laboratory of Electronics Technical Report No. 209, April, 1952.

[13] "A Method of Evaluating Aircraft Stability Parameters From Flight Test Data," by L. E. Schumacher, A. F. Technical Report WADC-TR-52-71, June, 1952.

Design of a Self-Optimizing Control System

By R. E. KALMAN,[1] NEW YORK, N. Y.

This paper examines the problem of building a machine which adjusts itself automatically to control an arbitrary dynamic process. The design of a small computer which acts as such a machine is presented in detail. A complete set of equations describing the machine is derived and listed; engineering features of the computer are discussed briefly. This machine represents a new concept in the development of automatic control systems. It should find widespread application in the automation of complex systems such as aircraft or chemical processes, where present methods would be too expensive or time-consuming to apply.

Introduction

THE art of the design of systems for the automatic control of dynamic processes of many different kinds (such as airplanes, chemical plants, military-weapon systems, and so on) has been reduced gradually to standard engineering practice during the years following World War II. In the simplest possible setting, the problem that the engineer faces in designing such automatic control systems is shown in Fig. 1. It is desired that the output of the process $c(t)$, which may be position, speed, temperature, pressure, flow rate, or the like, be as close as possible at all times to an arbitrarily given input $r(t)$ to the system. In other words, at all instants of time it is desired to keep the error $e(t) = r(t) - c(t)$ as small as possible. Control is accomplished by varying some physical quantity $m(t)$, called the control effort, which affects the output of the process.

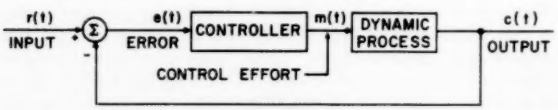


FIG. 1 BLOCK DIAGRAM OF SIMPLEST CONTROL PROBLEM

As long as the deviations from an equilibrium value of $r(t)$, $c(t)$, and therefore of $e(t)$ and $m(t)$, are small, the system can be regarded as approximately linear and there is a wealth of theoretical as well as practical information on which engineering design may be based. (When the system is not linear, present-day knowledge supplies only fragmentary suggestions for design; however, nonlinear effects are frequently of secondary importance.) It is generally agreed that the design of high-performance control systems is essentially a problem of matching the dynamic characteristics of a process by those of the controller. Practically speaking, this means that if the dynamic characteristics of the process are known with sufficient accuracy, then the characteristics of a controller necessary to give a certain desired type of performance can be specified. Usually, this amounts to writing down in quantitative terms the differential equations of the controller. Thus the design procedure can be divided roughly into the following distinct stages:

I Measure the dynamic characteristics of the process.
II Specify the desired characteristics of the controller.
III Put together a controller using standard elements (amplifiers, integrators, summers, electric networks, and so on) which has the required dynamic characteristics.

This subdivision of effort in designing a control system is oversimplified, but it will be a convenient starting point for the following discussion.

It has been pointed out by Bergen and Ragazzini (1)[2] that if a high degree of flexibility is desired in design stage (III), it is advantageous to use a sampled-data system. In principle, a sampled-data system is one where the controller is a digital computer. It is probably no exaggeration to say that, because of the great inherent flexibility of a digital computer, *any* desired controller characteristic is practically realizable. The use of a digital computer for the controller reduces stage (III) to a straightforward operation, like that of transcribing a handwritten manuscript by means of a typewriter.

Since the theory of linear control systems is well developed, stages (I–II) also can be made to consist of more-or-less standard procedures. Quick and convenient design even in stage (III) demands or at least suggests a digital computer; so the question arises whether or not stages (I–II) also can be reduced to completely mechanical operations which can be performed by a digital computer. Accordingly, the problem considered in this paper can be stated as follows:

To design a machine which, when inserted in the place of the controller in Fig. 1, will automatically perform steps (I–III), and set itself up as a controller which is optimum in some sense. The design of this machine is to be based on broad principles only. Its operation should require no direct human intervention but merely the measurements of $r(t)$ and $c(t)$.

In other words, such a machine, if it can be built, eliminates the lengthy, tedious, and costly procedure of engineering design—it is only necessary to connect the machine to *any* process. Thus the machine would seemingly eliminate the need for the control-systems engineer, but the latter can be reassured by the fact that the design of the machine itself is a far more ambitious and challenging undertaking than that of conventional control systems.

An even more decisive advantage of the machine over present-day design procedures is the following: In carrying out steps (I–III) it is generally taken for granted that the dynamic characteristics of the process will change only slightly under any operating conditions encountered during the lifetime of the control system. Such *slight* changes are foreseen and are usually counteracted by using feedback. Should the changes become large, the control equipment as originally designed may fail to meet performance specifications. Instances where difficulties of this type are encountered are:

(*a*) Changes of aircraft characteristics with speed.
(*b*) Chemical processes.
(*c*) Any large-scale control operation, where the nature of the system can be affected by uncontrolled and unforeseen factors.

By contrast, the machine can *repeat* steps (I–III) continually and thereby detect and make corrections in accordance with any

[1] Department of Electrical Engineering and Electronics Research Laboratories, Columbia University; formerly, Engineering Research Laboratory, E. I. du Pont de Nemours & Company, Wilmington, Del.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 14, 1957. Paper No. 57—IRD-12.

[2] Numbers in parentheses refer to the References at the end of the paper.

changes in the dynamic characteristics of a process which it controls. Such a control system operates always at or near some "optimum," provided only that changes in the dynamic characteristics of the controlled process do not occur very abruptly. It may be said that the machine adapts itself to changes in its surroundings—this may be regarded as an extension of the principle of feedback. The author prefers to call this property of the machine "self-optimization." The word "ultrastability" has been suggested also in a similar context by Ashby (2).

In the stated degree of generality, the problem is certainly not at a stage at present where any clear-cut ("unique") solution can be expected. Therefore this paper does not treat the general problem but presents a specific approach which leads to a practically satisfactory solution. This point is of considerable interest, since some earlier speculations relating to the problem were mostly of theoretical nature, without an attempt to appraise the difficulties (cost, complexity, and so on) of practical implementation (2–5). A machine based on the principles discussed in what follows actually has been built and will be described briefly in a later section.

It should be emphasized that the machine has been designed from a practical engineering point of view, rather than deduced from some law of physics or mathematics. The various single elements in the design of the machine are based on known principles. The choice between alternate possibilities in each stage of the design has been guided by efficiency and cost considerations. It is claimed that the over-all design uniting these principles in one machine is new and represents a major advance in regard to practicality over suggestions contained in the current literature.

General Design Considerations

From the technological point of view, it is clear that the machine discussed in the preceding section must be a computer. There are two possible choices, analog or digital computer. The latter choice is preferable. The reason is this. An analog computer is basically a method of simulating simple dynamic processes as they occur in the physical universe. The machine in question is required to simulate the actions of man, not of nature. This requires much greater flexibility and at the present state of computer technology such flexibility is provided only by digital computers.

The words "digital" and "analog" used here refer to the *external* characteristics of computers. Mathematically speaking, an analog computer performs the operations of analysis, such as differentiation, integration, computing logarithms, and so on, while a digital computer performs only arithmetic operations; namely, addition and multiplication. An analog computer operates on continuous functions (of time), the digital computer deals with discrete numbers. As far as the *internal* construction of these machines is concerned, it may happen that a computer which is called analog by its user contains discrete components (such as very fast counting circuits); and a computer which is called digital by its user may contain continuous components (such as potentiometers). Following these remarks, the computer that is described later may be called externally digital, internally analog.

In a digital computer, mathematical operations must be expressed (using approximations of various types) in numerical form. For instance, a function such as e^x must be computed by means of a series, which involves only repeated addition and multiplication. Another example is measuring the dynamic characteristics (transfer function or impulse response) of a process. Mathematically, this leads to the problem of solving an integral equation for which no satisfactory analog computing technique exists at present. On a digital computer the problem reduces to solving a set of simultaneous algebraic equations which is much simpler than solving an integral equation.

These considerations suggest the first fundamental design requirement:

(*A*) *The machine must be a digital computer.*

Recall now that the machine has a twofold job; namely, design and control. (i) It must measure the dynamic characteristics of the process and then determine the best form of the controller. (ii) It must control the process by providing the required control action $m(t)$. It is naturally desirable to keep these distinct functions independent. Therefore:

(*B*) *The operations necessary for designing a suitable controller must not be allowed to interact with the control action itself.*

It will be seen later that this requirement cannot be satisfied completely; the degree to which it must be relaxed to provide satisfactory operation is one of the unanswered questions at present.

Special Design Considerations

There are several practical requirements, all quite self-evident, which must be satisfied if the machine is to fulfill the expectations presented in the Introduction. All of these are related to design problem (I).

The functioning of the machine must not be critically dependent on obtaining measurements with high accuracy. Determination of the dynamic characteristics of the process is based on knowledge of $m(t)$ and $c(t)$. Since the first of these is actually produced by the machine itself, it may be assumed to be known with arbitrary accuracy; $c(t)$, however, corresponds to some physical quantity such as temperature, flow, and so on, whose determination is always accompanied by errors due to the imperfect operation of measuring equipment. These errors are called *measurement noise*. The standard method of reducing measurement noise is to take a large number of measurements. This leads to the requirement:

(*C*) *The determination of the dynamic characteristics of the process must be based on a large number of measurements so as to minimize the effects of measurement noise.*

As pointed out in the Introduction, one of the potential advantages of such a machine is that it can constantly repeat the entire design procedure and thereby adjust itself in a manner corresponding to any changes in process characteristics. But because of requirement (*C*), the determination of process characteristics requires a large number of measurements, taking a (possibly) long period of time. Since the system characteristics at the end of a series of measurements may be appreciably different from what they were at the beginning of the series of measurements, it is clear that older measurements ("obsolete data") should not be regarded as being as good as more recent measurements. This may be stated as:

(*D*) *Among any two measurements of $c(t)$, the more recent one should be given the higher weight: Measurements of $c(t)$ made infinitely long ago should be given zero weight.*

The cost, size, probability of breakdown, and so on of the machine is roughly proportional to the number of computations it has to perform per unit time. Therefore other things being equal, the number of computations should be as small as possible:

(*E*) *The methods of numerical computation to be used in the machine should be highly efficient.*

This last requirement will make it possible also to choose between alternative methods of computation.

Computation of Transfer Function From Measurements

Sampling. We now examine in detail the problem of measuring the dynamic characteristics of the process to be controlled. To do this, the functions $m(t)$ and $c(t)$ must be known. Since, ac-

cording to requirement (*A*), the machine is to be a digital computer, it is necessary to replace $m(t)$ and $c(t)$, which are continuously varying functions of time, by sequences of numbers which are discretely varying functions of time. This process is known as *sampling*. The most common way of doing this is to perform measurements periodically. Let the sampling instants be $t = kT$, $k = 0, 1, 2, \ldots$, where T is called the sampling period. Then sampling replaces $m(t)$ and $c(t)$ by the sequences of numbers

$$\begin{matrix} m(0), & m(T), & m(2T), \ldots, m(kT), \ldots \\ c(0), & c(T), & c(2T), \ldots, & c(kT), \ldots \end{matrix} \qquad k = 0, 1, \ldots . \quad [1]$$

In order to simplify the notation, we frequently will write $m_k = m(kT)$ and $c_k = c(kT)$ from now on. As a result of the sampling process, all experimental information about the functions $m(t)$ and $c(t)$ is contained in the numbers [1]. The sampling process is illustrated in Fig. 2.

The theory of linear control systems in which some of the controlled quantities are subject to sampling (the so-called sampled-data systems) is well developed. For further information, see Ragazzini and Zadeh (6) and Truxal (7).

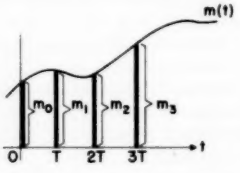


FIG. 2 SAMPLING PROCESS

Step Response of the Process. If the process is *linear, time-invariant, and stable*, it is well known that $c(t)$ is related to $m(t)$ by the convolution integral

$$c(t) = \int_{-\infty}^{t} h(t - u)dm(u) \ldots \ldots \ldots \ldots \quad [2]$$

where $h(t)$ is the step-function response of the process; $h(t) = 0$ when $t < 0$. Once $h(t)$ (or one of its equivalent forms, for instance, its Laplace transform) is known, the dynamic behavior of the process in question is completely characterized. But to find $h(t)$ given $m(t)$ and $c(t)$ by means of Equation [2] requires solving an integral equation which is a very difficult task.

If we consider now the closed-loop system shown in Fig. 1, it is clear that the input $m(t)$ to the process is the output of the self-optimizing controller. Therefore $m(t)$ must depend on the output of a digital computer; in other words, $m(t)$ must be a function of time which is completely determined by its values m_k at the sampling instants. To construct a function $m(t)$ from the series of numbers m_k which has a definite value at every instant of time calls for some method of interpolation. The simplest and practically most frequently used method (6, 7) is to hold the value of $m(t)$ constant after each sampling instant until the next sampling instant. In mathematical notation

$$m(t) = m_k, \qquad kT \leqslant t < (k + 1)T \ldots \ldots \ldots \quad [3]$$

Assuming that $m(t)$ is given by Equation [3], it is easy to show that the convolution integral Equation [2] reduces to the sum

$$c(t) = \sum_{l=-\infty}^{lT<t} h(t - lT)(m_l - m_{l-1}) \ldots \ldots \ldots \quad [4]$$

Noting that $h(kT) = 0$ for all $k < 0$, and considering only sampled values of $c(t)$ and $h(t)$, Equation [4] can be rewritten in the simpler form

$$c_k = \sum_{l=-\infty}^{l=k} (h_{k-l} - h_{k-l-1})m_l = \sum_{l=-\infty}^{l=k} g_{k-l}m_l \ldots \ldots \quad [5]$$

where the g_k's are recognized as the samples of the response of the system to a unit pulse. According to Equation [5], the dynamic behavior of the process is now represented by the sequence of numbers

$$g_0 = h(0), \quad g_1 = h(T) - h(0), \ldots,$$
$$g_k = h(kT) - h[(k - 1)T], \ldots$$

Moreover, if the input-output sequences [1] are known after some sampling instant, say, $k = 0$, then the numbers g_k can be determined by solving an infinite set of simultaneous linear algebraic equations given by Equation [5]. Since $h_k \to$ const with $k \to \infty$ (otherwise the process would not be stable and therefore Equation [5] would not be valid at all) it can be assumed in practice that $h_k = h_N$ for all $k > N$ if N is sufficiently large. This assumption means that $g_k = 0$ for all $k > N$ so that only a *finite* set of linear algebraic equations has to be solved to get the g_k.

But even with this simplification it would be quite inefficient to represent the process by means of the g_k because this would require a large amount of storage in the digital computer. For instance, if the step response of the process is

$$h(t) = 1 - \exp(-t/\tau)$$

$$g_0 = 0, \qquad g_k = [\exp(T/\tau) - 1] \exp(-kT/\tau), \qquad k \geq 1$$

then approximately $N = 5\tau/T$ numbers are necessary if the error due to neglecting the terms g_k, $k > N$ is to be less than 1 per cent. If fast control is required, the time constant of the closed-loop system must be much less than τ; on the other hand, the response of the closed-loop system on the average cannot take place in less than T seconds. Thus τ/T must be large, which means that a large number of values of g_k must be stored. This and other practical considerations to be discussed later indicate that the numbers g_k do not represent the dynamic characteristics of a process efficiently.

Pulse Transfer Function. A different way to represent a dynamic process is to assume that there is a linear differential equation relating $m(t)$ to $c(t)$. Consequently, m_k and c_k may be assumed to be related by means of a linear difference equation

$$c_k + b_1c_{k-1} + \ldots + b_nc_{k-n} = a_0m_k + a_1m_{k-1} + \ldots + a_qm_{k-q} \ldots \ldots \quad [6]$$

where the a_i and b_i are real constants and b_0 has been set arbitrarily equal to unity. If the differential equation relating $m(t)$ and $c(t)$ is known, the Difference Equation [6] can be derived readily using the theory of sampled-data systems. Such a derivation shows that in general $q = n$. By rearranging Equation [6], it follows that c_k can be expressed in terms of previous inputs and outputs

$$c_k = a_0m_k + a_1m_{k-1} + \ldots + a_nm_{k-n} - b_1c_{k-1} - \ldots - b_nc_{k-n} \ldots . \quad [6a]$$

Usually $a_0 = 0$, since most physical systems do not respond instantaneously. The theoretical difference between Equations [6a] and [4] is that in the latter case in principle all past inputs are needed to determine the present output while in the former case only a finite number of past inputs and outputs is needed. The practical difference is that when the system is known to be governed by a difference equation, much fewer a_i and b_i than g_k are needed to represent the system.

Using the notation $z^ic_k = c_{k+i}$ (where i is any integer), it is possible to write down the following basic relationship between the g_k defined by Equation [5] and the a_i and b_i defined by Equation [6]

$$G(z) = \frac{a_1 z^{-1} + \ldots + a_n z^{-n}}{1 + b_1 z^{-1} + \ldots + b_n z^{-n}} = g_1 z^{-1} + g_2 z^{-2} + \ldots + g_k z^{-k} + \ldots \quad [7]$$

where the right-hand term is obtained by the formal expansion of the rational fraction $G(z)$ by long division according to ascending powers of z^{-1}. The first term, g_0, is missing because it was assumed that $a_0 = 0$ which implies that $h_0 = g_0 = 0$. The function $G(z)$ is called the *pulse transfer function* of the process (6, 7). It has the same role in the analysis of linear sampled-data systems as the transfer function (Laplace transform of a differential equation) in the analysis of linear continuous systems.

The number of the a_i and b_i used to represent the process is based also on an assumption as to what the value of n should be. This is a matter of approximation; in other words, n should be chosen sufficiently large so that the a_i and b_i represent the process with some desired accuracy. But the characteristics of the process are not known in advance so that some initial guess must be made about n in setting up the machine. It is, of course, possible in principle to let the machine check the adequacy of this initial guess once experimental data about the process are available. For simplicity, however, the machine discussed in this paper was designed to operate with a fixed choice of n ($n = 2$).

Finally, it should be recalled that use of the numbers g_k is feasible only if the process is stable. No such restriction is inherent in the representation by Equation [6].

To summarize, the first step in the design of the machine is:

(i) *The dynamic characteristics of the process are to be represented in the form of Equation [6], the coefficients of which are to be computed from measurements. The number $n = q$ is assumed arbitrarily. In general, the higher n, the more accurate the representation of the process by the Difference Equation [6].*

Method of Determining Coefficients. According to design requirement (*C*), the coefficients in Equation [6] must be determined from a large number of measurements. This can be done as follows: Suppose we make a particular guess for the a_i and b_i at the Nth sampling instant. Let us denote these assumed values by $a_i(N)$ and $b_i(N)$, and compute all the past values of c_k using this particular set of coefficients and Equation [6a]. Denoting by $c_k^*(N)$ the values of the output computed in this way, we have

$$c_k^*(N) = -b_1(N)c_{k-1} - b_2(N)c_{k-2} - \ldots - b_n(N)c_{k-n} + a_1(N)m_{k-1} + a_2(N)m_{k-2} + \ldots + a_n(N)m_{k-n} \quad [8]$$

$$k = 0, 1, \ldots, N$$

A convenient measure of how good this choice of coefficients, in the light of past measured data, is the mean squared error

$$\frac{1}{N}\sum_{k=0}^{k=N} \epsilon_k^2(N) = \frac{1}{N}\sum_{k=0}^{k=N} [c_k - c_k^*(N)]^2 \quad [9]$$

where $\epsilon_k^2(N)$ represents the squared error between measured values c_k in the past and the predicted values $c_k^*(N)$ based on a certain choice of coefficients made at the Nth sampling instant; choosing the coefficients $a_i(N)$ and $b_i(N)$ in such a fashion that the mean squared error Equation [9] is a minimum called *least-squares filtering*. In general, any method for determining the $a_i(N)$ and $b_i(N)$ differs from least-squares filtering only in the form of the appropriate expression to be minimized. The advantage of least-squares filtering is that the computations can be carried out fairly simply (see Appendix), which is usually not the case if other types of error expression are used.

In view of design requirement (*D*), the more recent measurements should receive greater weight than very old ones, since the process dynamics may change with time. To meet this requirement, we proceed as follows: Let $W(t)$ be a continuous, monotonically decreasing function of time such that

$$\left.\begin{aligned} W(0) &= 1 \\ 0 < W(t) < 1,\ 0 < t &< \infty \\ W(\infty) &= 0 \\ \int_0^\infty W(t)dt &< \infty \end{aligned}\right\} \quad [10]$$

A function satisfying such conditions is called a *weighting function*. Writing W_k for $W(kT)$, the final criterion of determining the coefficients may be stated as follows: Choose $a_i(N)$, $b_i(N)$ in such a way that the expression

$$E(N) = \sum_{k=0}^{k=N} \epsilon_k^2(N) W_{N-k} \quad [11]$$

is a minimum. In other words the errors which would have been committed with the present choice of the coefficients $N - k$ sampling periods ago are to be weighted by a number $0 < W_{N-k} < 1$. Practically speaking, this means that the coefficients are calculated by disregarding errors which would have been committed in predicting the output a very long time ago (when the process may have been different) but trying to keep errors in predicting recent outputs small. None of these considerations, however, determines the precise form of the function $W(t)$; this question will be settled later so that an efficient computation procedure is obtained. We now state the second step in the design of the machine:

(ii) *The coefficients a_i and b_i should be determined anew at each sampling instant so as to minimize the weighted mean-square error $E(N)$.*

Numerical Solution of Weighted Least-Squares Filtering Problem. The explicit process necessary to determine the $a_i(N)$ and $b_i(N)$ requires, even after numerous simplifications, lengthy and somewhat involved calculations. These are discussed and recorded in detail in the Appendix. Only a few remarks are given here:

1 It is necessary to compute a number of so-called pseudo-correlation functions in order to write the error expression $E(N)$ in a simple form. These pseudo-correlation functions embody all measurement data up to the Nth sampling instant which is necessary to compute $E(N)$. To compute $E(N + 1)$, it is necessary to modify the pseudo-correlation functions so as to include the data received at the $(N + 1)$st sampling instant. It turns out that this process can be carried out in a simple way only if W_k is the unit pulse response (cf. Equations [5] and [7]) of a linear system governed by a difference equation. Then computation of the pseudo-correlation functions is carried out by passing products of measured values of m_k and c_k through a linear low-pass filter.

2 In order to apply Equation [6] to characterize a process, it is necessary that m_k and c_k be measured with respect to two reference values m_r and c_r such that, if m_r is a constant input to the system, c_r is the output in the steady state. Since the correct choice of such reference levels is not known in general, they must not enter into the computations of the type of Equation [6a]. In practice, the reference levels are usually determined by extraneous considerations such as calibration and range of measuring instruments. One way of avoiding the effect of incorrect reference levels (so-called *bias errors*) is to pass m_k and c_k through identical high-pass filters. After a sufficiently long period of time the bias errors, which are equivalent to a constant input to the filter, will be attenuated by an arbitrarily large factor at the output of an appropriately designed high-pass filter.

After the pseudo-correlation functions have been obtained, the determination of the coefficients reduces to solving a set of

simultaneous linear algebraic equations. To do this efficiently, an iteration procedure is used; it turns out that high-pass filtering m_k and c_k (which is equivalent approximately to subtracting the instantaneous mean value of these series of numbers) is a necessary requirement to insure the convergence of the iteration procedure.

The third step in the design is as follows:

(iii) *The calculations necessary for determining the coefficients consist of modifications of the classical least-squares filtering procedure and are given in the Appendix.*

Optimal Adjustment of Controller

Once the pulse-transfer function of the process to be controlled has been obtained, the synthesis of an "optimal" controller as a set of difference equations becomes a routine task (1, 8, 9).

It is not easy to agree, however, on what constitutes optimal control. The design of an optimal controller depends in general on two considerations:

(*a*) The nature of the input and disturbance signals to the system.

(*b*) The performance criterion used.

For instance, the inputs to the system may consist of step functions of various magnitudes; the performance criterion may be the length of time after the application of the step required by the control system to bring the error within prescribed limits. Or the input may consist of signals which are defined only in the statistical sense, in which case a reasonable performance criterion is the mean squared value of the error signal.

To include in the design of the machine means by which the machine can decide what class of input signals it is subjected to and what type of optimal controller should be used appears to be too ambitious a task at the present time. For this reason, in the practical realization of the machine (see the section Description of Computer), a prearranged method of optimizing the controller was used.

This method was described in a recent note by the author (8). The input signals are to consist of steps. The controller is to be designed in such a fashion that the error resulting from a step input becomes zero in minimum time and remains zero at all values of time thereafter. As a result of these assumptions the optimal controller is described by a difference equation whose coefficients are simple multiples of the coefficients of the pulse transfer function (see Equation [25] in the Appendix.)

We note the last step in the design:

(iv) *The choice of an optimal controller is largely arbitrary, depending on what aspect of system response is to be optimized. The determination of the coefficients in the describing equations of the controller is a routine matter if the coefficients of the pulse-transfer function are known.*

Summary of Machine Organization

Since the describing equations of the self-optimizing controller are somewhat involved, it is helpful to visualize the various computation processes as shown in Fig. 3.

Numbers in brackets indicate equations which characterize the particular operations performed. It should be remembered, of course, that there are many pseudo-correlation functions, coefficients, and so on, to be computed, some of which are indicated only in a schematic fashion.

It is perhaps worth while to emphasize that the closed-loop system consisting of the self-optimizing machine and the process is highly nonlinear. The principal nonlinear operations are:

(*a*) The multiplications before the input to low-pass filters whose outputs are the pseudo-correlation functions.

(*b*) The determination of controller coefficients.

These nonlinear operations have made it necessary to design the self-optimizing machine step by step. There exists at present no general theory for the design of nonlinear control systems of this type.

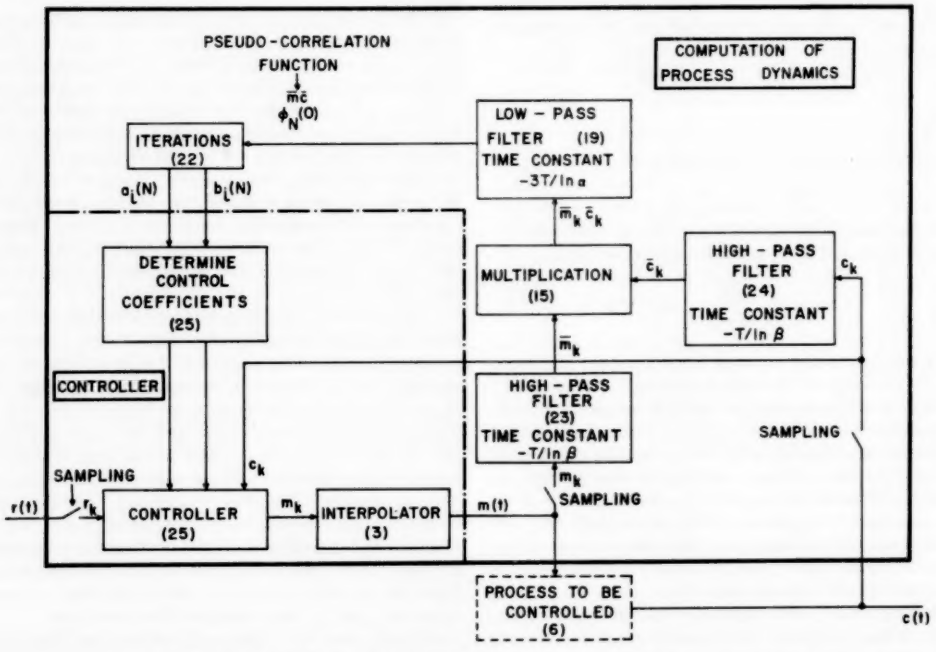


Fig. 3 Block Diagram of Computation Steps for Self-Optimizing Controller

Unsolved Questions

According to the preceding discussion, the operation of the self-optimizing system depends mainly on the accuracy of the computation of the pulse-transfer function from measurement data. Now suppose that the system is under very good control and that the input and disturbances to the system are nearly constant. In that case m_k and c_k will vary only very slightly about their equilibrium values. As a result, the numbers $\bar{m}_k$ and $\bar{c}_k$ (approximately the deviations of m_k and c_k from equilibrium) which are the inputs to the computation process determining the transfer function will be small and of roughly the same order of magnitude as the measurement noise. Under such circumstances, the transfer function cannot be computed very accurately. If the transfer function is not known accurately, then the controller cannot be set up accurately either and the system will not be operating optimally. But then the control will be less good and the deviations from the equilibrium values will increase. This, in turn, will improve the signal-to-noise ratio of the quantities $\bar{m}_k$ and $\bar{c}_k$; the computation of the transfer function will be more accurate, control action more nearly optimal, and so on. This shows that the operation of the system is limited basically by measurement noise. The fluctuations around the equilibrium condition must always be large enough to measure the transfer function with reasonable accuracy even in face of measurement noise. Thus the operation of the system depends on not being entirely at rest; if it were, it is impossible to say anything about the dynamic characteristics of the controlled process. A more precise answer to the problem involved here calls for further study.

Let us now examine qualitatively the effect of the choice α and β (cf. Fig. 3 and Appendix, Equations [19, 23, 24]) on this aspect of system performance. If α is very close to unity, the computation of the pulse-transfer function involves a large number of samples of $\bar{m}_k$ and $\bar{c}_k$ so that even if the system is at rest, i.e., $\bar{m}_k$ and $\bar{c}_k$ are practically zero, the computation of the pulse-transfer function is not affected for a long time, because the system "remembers" results of old measurements. On the other hand, if the process dynamics change rapidly in time, then α should be chosen fairly small because otherwise the computed transfer function will not be the actual transfer function. Thus α is a design parameter whose choice depends somewhat on the nature of a particular situation encountered. There is no reason, of course, why the system cannot adjust α also, but this is a problem beyond the scope of this paper.

The choice of β is guided by similar considerations. If the inputs to the system change slowly then β should be very close to unity for then the low-frequency components in m_k and c_k (slow "drift" about equilibrium point) will be very heavily attenuated. If the system is a more lively one, i.e., m_k and c_k fluctuate appreciably in time due to the effect of inputs or disturbances acting on the system, the β should be chosen smaller to improve the transient response of the high-pass filter. Thus β is another design parameter for the self-optimizing system.

Additional possibilities for improving these aspects of system operation should be considered in future work. More complicated weighting-functions and high-pass filters, suspending the operation of transfer-function computation when signal-to-noise levels become too low, putting in periodic test signals to check the operation of various parts of the computer, and the like, are some topics for future research.

Description of Computer

As soon as the operations discussed in the foregoing sections have been reduced to a set of numerical calculations (see Appendix) the machine has been synthesized in principle. This means that any general-purpose digital computer can be programmed to act as the self-optimizing machine.

In practical applications, however, a general-purpose digital computer is an expensive, bulky, extremely complex, and somewhat awkward piece of equipment. Moreover, the computational capabilities (speed, storage capacity, accuracy) of even the smaller commercially available general-purpose digital computers are considerably in excess of what is demanded in performing the computations listed in the Appendix.

For these reasons, a small special-purpose computer was constructed which could be called externally digital and internally analog according to the terminology in the section General Design Considerations. Briefly, this computer is organized as follows:

The computer operates on numbers whose absolute values do not exceed unity. Each number is represented by a 60-cycle-per-sec (cps) voltage. Numbers are stored on multiturn potentiometers, by positioning a given potentiometer by means of a servo arrangement in such a fashion that its output voltage (with unit excitation) is a 60-cps signal of the required magnitude and sign. Numbers are added by feeding corresponding voltages into electronic summing circuits. Two numbers a and b are multiplied by the following well-known method: If output of the potentiometer with unit excitation is b, then the output of the potentiometer with excitation a will be ab. The storage locations and summers can be interconnected in such a fashion that, in any one step of computation, the computer is capable of performing any one of the following types of operations

$$\left.\begin{aligned} & a_1b_1 + a_2b_2 + \ldots + a_7b_7 = x \\ \text{or} \\ & a_1b_1c_1d_1 + a_2b_2c_2 + a_3b_3c_3 = x \\ \text{or} \\ & a_1b_1c_1d_1e_1f_1g_1h_1 = x \end{aligned}\right\} \quad \ldots\ldots\ldots [12]$$

and so on

where each quantity appearing on the left-hand side of Equations [12] is an arbitrary number; x is the desired result of the computation. The fact that several additions and multiplications can be performed simultaneously is very convenient from the standpoint of programming the computer. Usually, each of Equations [12] must be broken up into several parts in programming them on a general-purpose computer.

The front view of the computer, which is roughly of the size of an average filing cabinet, is shown in Fig. 4. Only connections for input-output signals appear on the front panel. The programming of the computer is achieved by inserting wires into a "patch panel" on top of the computer which is shown in Fig. 5. Almost every signal voltage inside the computer is brought out to some contact on the patch panel. This arrangement makes it possible to interconnect the basic components of the computer in any manner desired and also facilitates troubleshooting and maintenance. The disadvantage of a patch-panel type of programming is that the change of program is a time-consuming operation; however, this is of minor significance since the machine is intended to operate with a fixed program in any typical application. The control panel shown in Fig. 5 also contains means for changing the sampling rate and reading numbers into any one of the storage locations in the computer.

The wiring necessary to connect computer components with the patch panel, together with associated relays, timing and checking circuits takes up approximately one third of the volume of the computer. Another one third of the volume is required for the electronic circuits performing summation and multiplication and

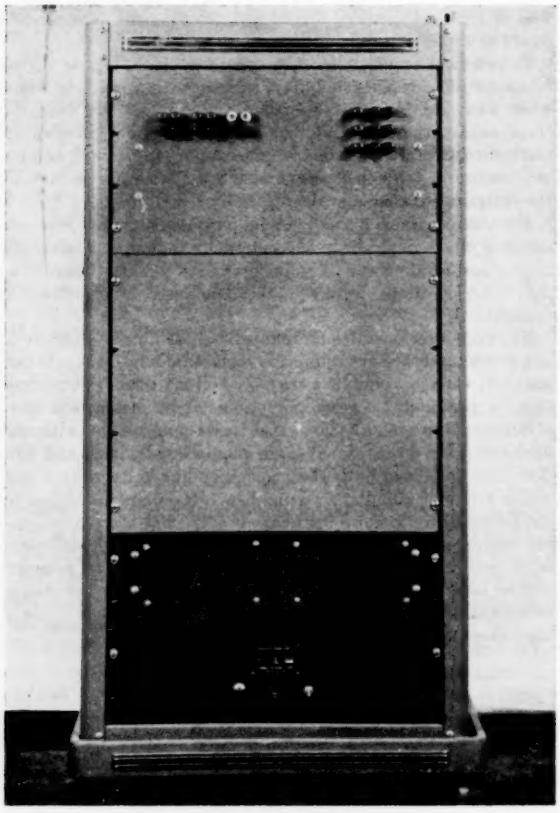


FIG. 4 FRONT VIEW OF COMPUTER

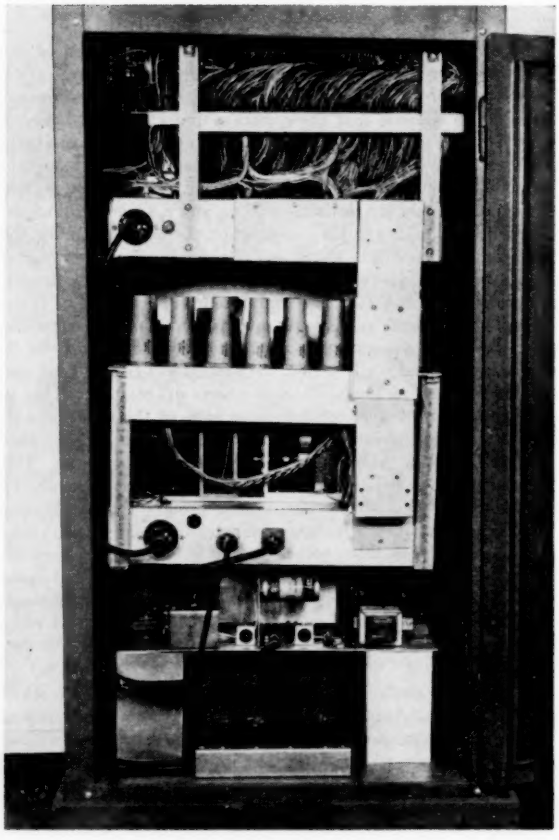


FIG. 6 REAR VIEW OF COMPUTER

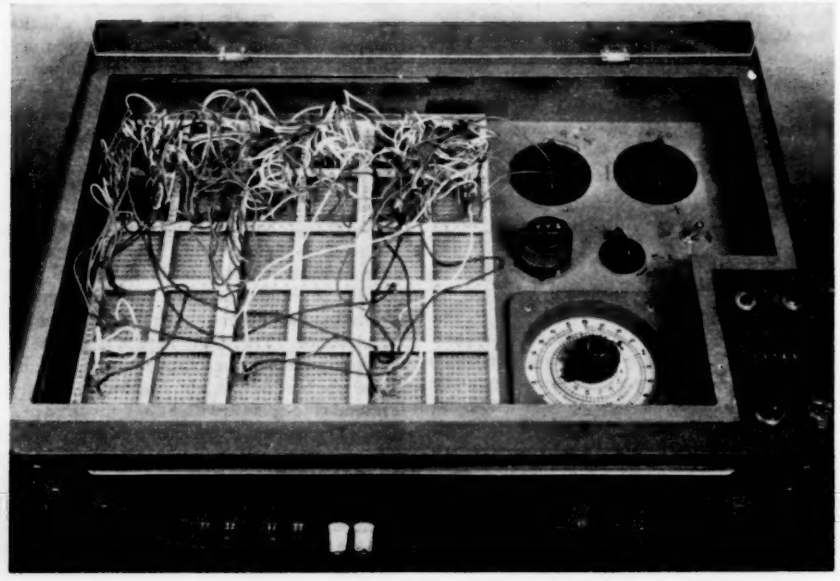


FIG. 5 CONTROL PANEL OF COMPUTER

the storage potentiometers. The remaining one third of space is taken up by power supplies. The internal arrangement of the computer is shown in the rear view of Fig. 6.

The computer described shows that the practical realization of a self-optimizing machine is well within the technological means available at the present time. Actually, the computer described was constructed in 1954/1955. The computer also represents savings in cost and complexity over currently available general purpose digital computers. On the other hand, when self-optimizing control of a large-scale installation is desired, in other words, when there are several dynamic processes to be controlled simultaneously and possibly in an interdependent fashion, then the general-purpose digital computer is much better matched to the problem both in terms of cost and computational capability.

Conclusions

This paper shows the feasibility of mechanizing much of the process by which automatic control systems for standard applications are being designed today. The amount of numerical computations necessary for accomplishing this is relatively modest (after the numerous simplifications discussed) and can be readily implemented in practice at moderate cost.

More importantly, however, the machine described here is an ideal controller since it needs merely to be interconnected with the process to be controlled to achieve optimum control after a short transitory period and hold it thereafter even if the process characteristics change with time. The task of the control engineer of the future will be not to design a specific system, but to improve the principles on which machines of the type described here will operate. Unlike his predecessor, the stock in trade of the new control-systems engineer will not be the graph paper, the slide rule, or even the analog computer but a firm and deep-seated understanding of the fundamental principles, physical and mathematical, on which automatic control is based. The drudgery of computing will be taken over by machines but the challenge of thinking remains.

Acknowledgments

The research reported here was supported by the Engineering Research Laboratory, E. I. du Pont de Nemours & Co., Wilmington, Del., to whom the author is indebted for permission to publish this paper. The author wishes also to thank various members of the Engineering Research Laboratory for their help and interest during the progress of this work, and to Dr. J. R. Ragazzini, Columbia University, for several stimulating discussions.

References

1 "Sampled-Data Processing Techniques for Feedback Control Systems," by A. R. Bergen and J. R. Ragazzini, Trans. AIEE, vol. 73, part II, 1954, pp. 236–247.

2 "Design for a Brain," by W. R. Ashby, John Wiley & Sons, Inc., New York, N. Y., 1952.

3 "Possibilities of a Two Time Scale Computing System for Control and Simulation of Dynamic Systems," by H. Ziebolz and H. M. Paynter, Proceedings of the National Electronics Conference, vol. 9, 1953, pp. 215–223.

4 "Determination of System Characteristics From Normal Operating Records," by T. P. Goodman and J. B. Reswick, Trans. ASME, vol. 77, 1955, pp. 259–268.

5 "Self-Optimizing Systems," by E. G. C. Burt, preprint for International Control Systems Conference, Heidelberg, Germany, September, 1956.

6 "The Analysis of Sampled-Data Systems," by J. R. Ragazzini and L. A. Zadeh, Trans. AIEE, vol. 71, part II, 1952, pp. 225–234.

7 "Automatic Feedback Control System Synthesis," by J. G. Truxal, McGraw-Hill Book Company, Inc., New York, N. Y., 1955.

8 R. E. Kalman, discussion of reference (1), Trans. AIEE, vol. 73, part II, 1954, pp. 245–246.

9 "Digital Controllers for Sampled-Data System," by J. E. Bertram, Trans. AIEE, vol. 75, part II, 1956, pp. 151–159.

10 "Introduction to Numerical Analysis," by F. B. Hildebrand, McGraw-Hill Book Company, Inc., New York, N. Y., 1956.

11 "Numerical Analysis," by W. E. Milne, Princeton University Press, Princeton, N. J., 1949.

Appendix

The following is the detailed derivation of the complete set of equations characterizing the self-optimizing controller in the special case when $n = 2$ in the Difference Equation [6]. Using these equations, any digital computer may be programmed to act as a self-optimizing controller. When $n > 2$, the required equations can be obtained similarly.

First of all, instead of performing the computations required to minimize Equation [11] at every sampling instant, they may be performed at every qth (where q is a positive integer) sampling instant. This does not affect the reasoning in the section Method of Determining Coefficients, and results in considerable simplification in the required computations. With this change, the error expression Equation [11] becomes

$$E(N) = \sum_{j=0}^{j=N/q} \epsilon_{qj}^2(N) W_{N-qj} \quad \ldots\ldots\ldots\ldots [13]$$

where $k = qj$ and N is a number divisible by q.

Now assume that $n = 2$ in Equation [6]. Using the recurrence relation Equation [8], $\epsilon_{qj}^2(N)$ can be written as

$$\left.\begin{aligned}
\epsilon_{qj}^2(N) &= [c_{qj} - c_{qj}^*(N)]^2 \\
&= c_{qj}^2 + b_1^2(N)c_{qj-1}^2 + b_2^2(N)c_{qj-2}^2 \\
&+ 2b_1(N)c_{qj}c_{qj-1} + 2b_2(N)c_{qj}c_{qj-2} \\
&\qquad + 2b_1(N)b_2(N)c_{qj-1}c_{qj-2} \\
&- 2a_1(N)c_{qj}m_{qj-1} - 2a_2(N)c_{qj}m_{qj-2} \\
&- 2b_1(N)a_1(N)c_{qj-1}m_{qj-1} \\
&\qquad - 2b_1(N)a_2(N)c_{qj-1}m_{qj-2} \\
&- 2b_2(N)a_1(N)c_{qj-2}m_{qj-1} \\
&\qquad - 2b_2(N)a_2(N)c_{qj-2}m_{qj-2} \\
&+ a_1^2(N)m_{qj-1}^2 + a_2^2(N)m_{qj-2}^2 \\
&\qquad + 2a_1(N)a_2(N)m_{qj-1}m_{qj-2}
\end{aligned}\right\} \quad \ldots [14]$$

The measured values of c and m occur in Equation [14] always in terms of the type

$$c_{qj-r}c_{qj-s} \qquad c_{qj-r}m_{qj-s} \qquad m_{qj-r}m_{qj-s} \quad \ldots\ldots\ldots [15]$$

where $r, s = 0, 1, 2$. If we now let

$$q = n + 1 = 3$$

then it is clear that factors of the same type will be multiplied by the same coefficients in Equation [14], regardless of the value of j. This property does not arise when $q < 3$. Using the symmetry introduced by the particular choice of q, $E(N)$ can be put in a simpler form by defining the *pseudo-correlation functions*

$$\left.\begin{aligned}
\phi_{N-r}^{cc}(r-s) &= \sum_{j=1}^{j=N/3} c_{3j-r}c_{3j-s}W_{N-3j} \\
\phi_{N-r}^{cm}(r-s) &= \sum_{j=1}^{j=N/3} c_{3j-r}m_{3j-s}W_{N-3j} \\
\phi_{N-r}^{mm}(r-s) &= \sum_{j=1}^{j=N/3} m_{3j-r}m_{3j-s}W_{N-3j}
\end{aligned}\right\} \quad \ldots\ldots [16]$$

With these definitions, $E(N)$ can be written as follows, arranging the terms in the same fashion as in Equation [14]

$$\left.\begin{aligned} E(N) = {} & \phi_N{}^{cc}(0) + b_1{}^2(N)\phi_{N-1}{}^{cc}(0) + b_2{}^2(N)\phi_{N-2}{}^{cc}(0) \\ & + 2b_1(N)\phi_N{}^{cc}(-1) + 2b_2(N)\phi_N{}^{cc}(-2) \\ & \qquad + 2b_1(N)b_2(N)\phi_{N-1}{}^{cc}(-1) \\ & - 2a_1(N)\phi_N{}^{cm}(-1) - 2a_2(N)\phi_N{}^{cm}(-2) \\ & - 2b_1(N)a_1(N)\phi_{N-1}{}^{cm}(0) \\ & \qquad - 2b_1(N)a_2(N)\phi_{N-1}{}^{cm}(-1) \\ & - 2b_2(N)a_1(N)\phi_{N-2}{}^{cm}(1) - 2b_2(N)a_2(N)\phi_{N-2}{}^{cm}(0) \\ & + a_1{}^2(N)\phi_{N-1}{}^{mm}(0) + a_2{}^2(N)\phi_{N-2}{}^{mm}(0) \\ & \qquad + 2a_1(N)a_2(N)\phi_{N-1}{}^{mm}(-1) \end{aligned}\right\} \quad \ldots [17]$$

Remark. The conventional definition of correlation functions is

$$\phi_N{}^{cc}(r) = \frac{1}{N}\sum_{k=0}^{k=N} c_k c_{k+r}$$

To evaluate this function iteratively, as is done in Equation [19] for pseudo-correlation functions, it would be necessary to compute

$$\phi_N{}^{cc}(r) = c_N c_{N+r}/N + (N-1)\phi_{N-1}{}^{cc}(r)/N$$

Since the factor $(N-1)/N$ cannot be calculated accurately enough as $N \to \infty$, such an iterative calculation would be impractical.

The pseudo-correlation functions can be evaluated iteratively as follows: Suppose that, in addition to meeting Conditions [10], the weighting function W_k is a sequence of numbers such as the g_k given by Equation [7]. Then it follows that the pseudo-correlation functions can be regarded as the output of a linear system governed by a difference equation, whose input consists of products such as Equation [15]. In particular, if we let

$$W^{sj} = \alpha^j \qquad (0 < \alpha < 1) \ldots\ldots\ldots\ldots [18]$$

then every pseudo-correlation function satisfies a first-order difference equation of the type

$$\phi_{sj-r}{}^{cm}(r-s) - \alpha\phi_{s(j-1)-r}{}^{cm}(r-s) = c_{sj-r}m_{sj-s} \ldots [19]$$

According to Equation [17] the determination of the coefficients of the pulse-transfer function requires first that all input-output data (the measured values of c and m) be consolidated into the pseudo-correlation functions. Because of the recurrence relation Equation [19], the computation of the latter is quite simple, since to get the pseudo-correlation functions at the Nth sampling instant requires only the knowledge of the same functions at the end of the $(N-3)$th sampling instant, plus the values of c_{N-2}, c_{N-1}, c_N, m_{N-2}, m_{N-1}. Once the new pseudo-correlation functions have been computed, the data measured during the preceding three sampling periods can be discarded and the system is ready to receive new data. Thus the use of the pseudo-correlation functions and the choice of a suitable weighting function greatly simplifies the implementation of mean-square filtering.

In order that $E(N)$ be a minimum with respect to the a_i and b_i, it is *necessary* that the partial derivatives

$$\frac{\partial E(N)}{\partial a_i} = 0 \qquad \frac{\partial E(N)}{\partial b_i} = 0 \quad (i = 1, 2, \ldots, n) \ldots [20]$$

vanish. The proof that these conditions are also *sufficient* to insure the existence of a minimum of $E(N)$ is quite difficult. Refer to Milne (11) for discussion of a closely related problem.

The Conditions [20] lead to four linear equations in the coefficients $a_1(N)$, $a_2(N)$, $b_1(N)$, $b_2(N)$ as follows

$$\left.\begin{aligned} & a_1(N)\phi_{N-1}{}^{mm}(0) + a_2(N)\phi_{N-1}{}^{mm}(-1) - b_1(N)\phi_{N-1}{}^{cm}(0) \\ & \qquad - b_2(N)\phi_{N-2}{}^{cm}(1) = \phi_N{}^{cm}(-1) \\ & a_1(N)\phi_{N-1}{}^{mm}(-1) + a_2(N)\phi_{N-2}{}^{mm}(0) \\ & \qquad - b_1(N)\phi_{N-1}{}^{cm}(-1) - b_2(N)\phi_{N-2}{}^{cm}(0) = \phi_N{}^{cm}(-2) \\ & -a_1(N)\phi_{N-1}{}^{cm}(0) - a_2(N)\phi_{N-1}{}^{cm}(-1) \\ & \qquad + b_1(N)\phi_{N-1}{}^{cc}(0) + b_2(N)\phi_{N-1}{}^{cc}(-1) = -\phi_N{}^{cc}(-1) \\ & -a_1(N)\phi_{N-2}{}^{cm}(1) - a_2(N)\phi_{N-2}{}^{cm}(0) \\ & \qquad + b_1(N)\phi_{N-1}{}^{cc}(-1) + b_2(N)\phi_{N-2}{}^{cc}(0) = -\phi^{cc}(-2) \end{aligned}\right\} \quad \ldots [21]$$

Any method for solving linear simultaneous equations can be used for finding the a_i and b_i from Equation [21]. However, the standard elimination methods (which, incidentally, are much more efficient than solving Equation [21] by Cramer's rule) require a rather large amount of storage and somewhat lengthy computations. These disadvantages become increasingly worse as n increases. However, an exact computation of a solution of Equation [21] is very wasteful in that, if a solution of Equation [21] at the $(N-3)$th sampling instant is available, then that solution is also an excellent guess for the solution of Equation [21] at the Nth sampling instant since the correlation function can have changed only slightly, unless a very small value of α is used. This suggests an *iteration* procedure for solving Equation [21], of which the simplest is the so-called Gauss-Seidel method (10).

Applying the Gauss-Seidel method to Equation [21] leads to the equations

$$a_1(N) = \frac{-a_2(N-3)\phi_{N-1}{}^{mm}(-1) + b_1(N-3)\phi_{N-1}{}^{cm}(0) + b_2(N-3)\phi_{N-2}{}^{cm}(1) + \phi_N{}^{cm}(-1)}{\phi_{N-1}{}^{mm}(0)} \ldots [22a]$$

$$a_2(N) = \frac{-a_1(N)\phi_{N-1}{}^{mm}(-1) + b_1(N-3)\phi_{N-1}{}^{cm}(-1) + b_2(N-3)\phi_{N-2}{}^{cm}(0) + \phi_N{}^{cm}(-2)}{\phi_{N-2}{}^{mm}(0)} \ldots [22b]$$

$$b_1(N) = \frac{a_1(N)\phi_{N-1}{}^{cm}(0) + a_2(N)\phi_{N-1}{}^{cm}(-1) - b_2(N-3)\phi_{N-1}{}^{cc}(-1) - \phi_N{}^{cc}(-1)}{\phi_{N-1}{}^{cc}(0)} \ldots [22c]$$

$$b_2(N) = \frac{a_1(N)\phi_{N-1}{}^{cm}(1) + a_2(N)\phi_{N-2}{}^{cm}(0) - b_1(N)\phi_{N-1}{}^{cc}(-1) - \phi_N{}^{cc}(-2)}{\phi_{N-2}{}^{cc}(0)} \ldots [22d]$$

If desired, the cycle of iterations just written down can be repeated to obtain better accuracy.

A necessary and sufficient condition for the convergence of the iteration Equations [22] is that the diagonal coefficients in Equations [21], i.e., $\phi_{N-1}{}^{mm}(0)$, $\phi_{N-2}{}^{mm}(0)$, $\phi_{N-1}{}^{cc}(0)$, $\phi_{N-2}{}^{cc}(0)$ should be larger in absolute value than any of the other coefficients in the same equation. To insure rapid convergence, it is highly desirable that the diagonal coefficients be as large as possible compared to the off-diagonal coefficients.

A glance at Equation [19] shows that the pseudo-correlation

functions just mentioned are always the sum of positive numbers because the right-hand side of Equation [19] is always positive, being a square. To make the pseudo-correlation functions corresponding to the off-diagonal elements in Equations [21] smaller in absolute value than the diagonal elements, the right-hand side of Equation [19] for these functions must be alternatively positive and negative. This can be achieved by subtracting from each c_k and m_k the average (mean) values of these quantities over a long period of time. Unless this is done, c_k and m_k might vary only slightly about a large average value in which case all the correlation functions will be approximately equal and the iteration Equation [22] will not converge fast enough, if at all.

To estimate the mean of a time series in a very reliable way is not an easy problem. In the present case, however, sophisticated statistical methods are not required because the precise knowledge of the mean is not important. The simplest procedure then is to put both c_k and m_k through identical high-pass filters which remove the slowly varying components (i.e., the mean) of these quantities. When the mean is constant in time, it is equal to the zero frequency component of the signal. The simplest high-pass filter on numerical data is represented by the difference equation

$$c_k - c_{k-1} = \bar{c}_k - \beta\bar{c}_{k-1} \qquad (0 < \beta < 1) \ldots\ldots [23]$$

where $\bar{c}_k$ is approximately equal to c_k − mean (c_k). The closer β is to 1, the better the removal of the mean if the latter is constant. On the other hand, if the mean varies β should be somewhat smaller for best results. A similar equation holds for $\bar{m}_k$

$$m_k - m_{k-1} = \bar{m}_k - \beta\bar{m}_{k-1} \qquad (0 < \beta < 1) \ldots\ldots [24]$$

A simple substitution in Equation [6] shows that $\bar{c}_k$ and $\bar{m}_k$ are related by the same difference equation as c_k and m_k. This is because if two quantities are linearly related, the relationship remains undisturbed if both quantities are put through identical linear filters. Thus the removal of the mean represented by Equations [23] and [24] does not affect the computation of the pulse-transfer function of the process to be controlled, except for greatly improving the convergence of the iteration process Equations [22]. Hence all pseudo-correlation functions should be computed using the $\bar{c}_k$ and $\bar{m}_k$.

It remains to show how the equations of the controller can be obtained from the knowledge of the coefficients of the pulse-transfer function. As mentioned earlier, the controller is to be digital. Using a method of synthesis due to the author (8), which yields the optimum design if the closed-loop system is to respond to a unit step input in minimal time without overshoot (for a given fixed sampling period T), the numbers necessary to specify the controller are very simply related to the coefficients of the pulse-transfer function of the process which is to be controlled. In fact, the difference equation specifying the controller is

$$[a_1(N) + a_2(N)]m_k - a_1(N)m_{k-1} - a_2(N)m_{k-2} = e_k + b_1(N)e_{k-1} + b_2(N)e_{k-2} \ldots\ldots [25]$$

where $$e_k = r_k - c_k$$

Equation [25] is valid for $N + 1 \leq k \leq N + 3$, after which a new set of coefficients must be used from the next determination of the pulse-transfer function. It should be noted that Equation [25] holds only if the (continuous) transfer function of the process is approximately $H(s) = K/(s + a)(s + b)$ with $a, b, K > 0$. If, for instance, $a = 0$, the form of Equation [25] is different. For methods of synthesizing digital controllers which are optimal in some other sense, see references (1, 9).

For convenience, the time sequence of computations to be performed during a cycle of $q = 3$ sampling periods is listed as follows.

$k = N - 2$

(1) Compute m_{N-2} using [25]
(2) Compute $\bar{c}_{N-2}$ using [23]
(3) Compute $\bar{m}_{N-2}$ using [24]
(4) Compute $\phi_{N-2}^{\bar{m}\bar{m}}(0)$, $\phi_{N-2}^{\bar{c}\bar{c}}(0)$, $\phi_{N-2}^{\bar{c}\bar{m}}(0)$ using Equation [19]

$k = N - 1$

(1) Compute m_{N-1} using [25]
(2) Compute $\bar{c}_{N-1}$ using [23]
(3) Compute $\bar{m}_{N-1}$ using [24]
(4) Compute $\phi_{N-1}^{\bar{m}\bar{m}}(0)$, $\phi_{N-1}^{\bar{m}\bar{m}}(-1)$, $\phi_{N-1}^{\bar{c}\bar{c}}(0)$, $\phi_{N-1}^{\bar{c}\bar{c}}(-1)$ $\phi_{N-1}^{\bar{c}\bar{m}}(0)$, $\phi_{N-1}^{\bar{c}\bar{m}}(-1)$, $\phi_{N-2}^{\bar{c}\bar{m}}(1)$ using Equation [19]

$k = N$

(1) Compute m_N using [25]
(2) Compute $\bar{c}_N$ using [23]
(3) Compute $\bar{m}_N$ using [24]
(4) Compute $\phi_N^{\bar{c}\bar{c}}(-1)$, $\phi_N^{\bar{c}\bar{c}}(-2)$, $\phi_N^{\bar{c}\bar{m}}(-1)$, $\phi_N^{\bar{c}\bar{m}}(-2)$ using Equation [19]
(5) Compute $a_1(N)$, $a_2(N)$, $b_1(N)$, $b_2(N)$ using [22]

Discussion

RANE L. CURL.[3] The author has presented with skill his proposal for a self-optimizing control system. He has also covered most of the limitations in both the theory and design of his machine. I will only mention perhaps one or two points that come to mind.

On the first stage of the author's procedure, *measure the dynamic characteristics of the process*, a difficulty would be met in most real processes of the regulatory type. The proposed method of determining the system characteristics is subject to error when the existence of an error signal is due to load disturbances entering between the *control effort* and the *output*. This error may be of two types. The first is from poor "response" information in the presence of noise, and is inherent in any method which does not use process response information over a very long time. The desire to make the self-optimizing machine respond to changes in process dynamics is anathema to obtaining a good measure of the transfer function in the presence of noise. The second type of error is inherent in *all* methods which determine process dynamics while the process is on closed loop control. The noise circulates in the loop and there exists a correlation between the *noise component* of the output $c(t)$, and the control effort $m(t)$.

The importance of the regulatory type of controller and the difficulty of obtaining good process dynamics when it is in use suggests a reason additional to that of the author as to why technological unemployment of control engineers will not result from this machine.

The author's use of $n = 2$, while a strict limitation, was, as the author correctly stated, a matter of convenience and not an inherent limitation. It would be of interest if the author would comment on the behavior of the machine described in his paper when used with systems having incompatible transfer functions, i.e., for processes for which Equation [25] does not represent the optimum controller.

The well known "optimizing" controller for adjusting a set point in order to maximize yield, profit, etc., introduces its own disturbance as a "tracer" on system performance. This is another possibility, in some cases, to computation suspension at low signal to noise ratios as in the author's machine.

I agree with the author that this machine does "represent. . . an advance. . . .in practicality over suggestions. . . .in the current literature." But I ask last the primary unanswered question: Does it work?

[3] Shell Development Company, Emeryville, Calif.

Author's Closure

Before taking up in detail the questions raised in Dr. Curl's discussion, the author wishes to answer his last and most important point, "Does it work?" The answer is, "Yes."

Dr. Curl's remarks on difficulties of determining the process transfer function amplify some of the matters discussed in the section, Unsolved Questions. As in any method of measurement based on statistical principles, the determination of the process dynamics depends on obtaining a large number of data with stationary statistical properties so that the effect of unwanted influences acting on the system can be averaged out. If the load disturbances have a nonzero mean value, then their effect on the plant appears as a shift in the operating point. The computation procedure determines the linear system dynamics for small deviations about this "phantom" operating point. Since the computation of the transfer function can take account of slow changes, shifts in the mean value of the load disturbances do not affect the operation of the system, provided that these shifts occur slowly relative to the sampling period. The accuracy of computation of the transfer function depends on the effective signal-to-noise ratio, that is, on the ratio of the mean-square value of the control effort $m(t)$ required under normal operating conditions to the mean-square value of the combined effect of load disturbances and measurement noise. When this ratio is too small, it may be improved by introducing special "test signals" into the plant, or the operation of the transfer-function computation may be temporarily suspended until the signal-to-noise ratio is improved.

The effect of circulating noise determines the maximum accuracy achievable by a self-optimizing system and can, in general, only be determined experimentally. If the effect is too large, more accurate instrumentation must be used. It should be borne in mind also that since the controller of a self-optimizing system is closely matched to the dynamics of the plant, any errors due to circulating noise can be rapidly corrected. In other words, measurement noise is not amplified by the system.

By way of illustration, it may be pointed out that measurements performed by the author using high-accuracy measuring equipment support the foregoing remarks. The computation of the transfer function of a crude third-order electrical analog (3 capacitors in cheap electronic circuity without voltage regulation) yielded the following experimental results, over about 500 sampling points:

Largest time constant $\tau_1 \cong$ constant $\pm$ 0.1 per cent
Next time constant $\tau_1/3 \cong$ constant $\pm$ 1.0 per cent
Smallest time constant $\tau_1/10 \cong$ constant $\pm$ 10 per cent

The high accuracy with which the dominant time constant τ_1 can be determined is quite remarkable. The variation is only slightly worse than the errors introduced by the measuring process. On the other hand, the large error in the determination of the smallest time constant is due to the combined effect of amplifier noise, temperature transients, and so forth. From this measurement, it may be concluded that the system may be regarded as effectively second-order. Indeed, the system could be controlled quite satisfactorily with a sampled-data controller with a fixed, second-order program. Conclusive results concerning the performance of the self-optimizing controller in an actual plant installation cannot be given here.

In conclusion, the author does not share Dr. Curl's pessimism that the presence of noise problems makes a self-optimizing system impractical. Probably the most serious practical difficulty barring better process control at the present time is the unavailability of accurate data on process dynamics. This difficulty can be circumvented in many cases by use of a self-optimizing controller. The author may not be unduly optimistic in expressing his feeling that (disregarding economic considerations) sufficient theoretical and technological know-how exists already to bring practical process control close to the best performance achievable in the light of the limitations imposed by physical measuring equipment.

Correlation Functions and Noise Patterns in Control Analysis

By HERMAN THAL-LARSEN,[1] BERKELEY, CALIF.

This paper describes the interplay of disturbance patterns and correlation functions when "noise" enters a control system, both in the fictitious case of the mathematical model and in the real process. In pointing out this interplay in the model and again in the real process, the application of correlation functions to control analysis is shown to have its limitations. At the same time the importance of discovering equivalent disturbance patterns is emphasized and a procedure for their discovery is developed.

Nomenclature

The following nomenclature is used in the paper:

a = process-output change, psi
c = controlled-variable change, psi
m = controller-output change, psi
n = disturbance, psi
r = controller set-point change, psi
s = Laplace-transform variable, 1/sec, here used as a convenient symbol for portraying transfer functions

Substitute $s = j\omega$ in frequency domain

$j = \sqrt{-1}$
ω = circular frequency, radians/sec
ω_{un} = undamped natural circular frequency, radians/sec
ζ = dimensionless damping ratio
α = attenuation rate, 1/sec
t = time, sec
τ = time base for $\phi(\tau)$, sec
T = time interval for convolution calculation
L = time interval used in definition of correlation function
$g(t), h(t)$ = unit-impulse responses
$\phi(\tau)$ or ϕ = correlation function, $(\text{psi})^2$
$\Phi(\omega)$ or Φ = spectral density, sec $(\text{psi})^2$

Note: Subscripts for ϕ and Φ indicate variables involved
k, p, i, g = summation limits

Introduction

Unwanted disturbances entering a control system make their presence known by what often appear to be random variations in the operating variables. Goodman and Reswick[2] show how correlation functions obtained from the normal operating records containing these random variations may be used to discover the dynamic characteristics of parts of a control system. Their method is appealing because, for purposes of analysis, it obviates the need of introducing additional disturbances such as step, impulse, or sinusoidal changes into a functioning plant. Consequently, when the dynamic characteristics of elements within a control loop must or should be measured, the good humor of the plant's operating staff need not be destroyed by the introduction of these additional disturbances. Even if these added upsets are tolerated by the plant's staff, the random variations already present within the control loop have a tendency to mask the true response.

Sometimes the dynamic characteristics of components can be measured without personnel annoyance only by using the statistical method. At other times, it is the only method that can be used.[3] Therefore, it is important that more information be made available concerning this approach and its application. The purpose of this paper is twofold: (1) To demonstrate the effect which the disturbance pattern has upon determination of system characteristics from correlation functions obtained from normal operating records, and (2) to show how the disturbance pattern may be discovered.

Procedure

A simple mathematical model of a control system was subjected to a purely random disturbance, to so-called "white noise." Various auto and crosscorrelation functions within and between the variables were then computed and plotted. As is known, white noise has an impulse-like autocorrelation function.

The same mathematical model was next subjected to a disturbance having an exponentially-decaying autocorrelation function. New correlation functions were computed and plotted which were compared to the first set obtained with white noise.

Then curves or patterns obtained from analyzing the behavior of the model under artificial disturbances were compared with curves calculated from data generated by a real process under a real disturbance pattern.

Finally, a method for discovering the unknown equivalent disturbance pattern by means of block-diagram inversion was developed.

Mathematical Model

Fig. 1 shows the block diagram for the mathematical model of a simple control system.

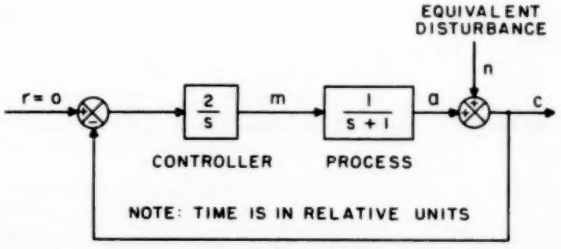


Fig. 1 Block Diagram of Simple Mathematical Model

[1] Lecturer in Mechanical Engineering, University of California. Mem. ASME.

[2] "Determination of System Characteristics From Normal Operating Records," by T. P. Goodman and J. B. Reswick, Trans. ASME, vol. 78, 1956, pp. 259–271.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-6.

[3] "The Application of an Analog Computer to the Measurement of Process Dynamics," by P. E. A. Cowley, ASME Paper No. 56—IRD-20.

In order to reduce computations to a minimum, the process block was simplified to a first-order system and the simplest controller equation which would still permit the system to oscillate was chosen. Constants were selected to produce a damping ratio of $\zeta = 0.35$, typical of many process-control systems.

Variables m, a, n, and c represent deviations from their mean value. The set point of the controller remained undisturbed, hence $r = 0$. Disturbance n represents the summation of all disturbances which enter the control loop between m and c to affect controlled variable c. Thus, n represents an "equivalent" disturbance.

Real Process

A block-diagram representation of the real process-control system is shown in Fig. 2. Here, pneumatic pressure is the controlled variable and the disturbance is produced by random fluctuations of pressure in the local water system. The water-pressure fluctuations were converted to pneumatic-pressure fluctuations by means of a Bourdon tube and a pneumatic nozzle-flapper amplifier.

To avoid overloading the control system, resultant pneumatic-pressure fluctuations were attenuated by means of a small throttling valve and tank. A Taylor Transet computing relay with adjustable suppression plus a volume booster allowed the disturbance to be injected finally into the control loop.

The mean value of the controlled pressure was set at 10 psig. The relay equation was, in psig

$$\text{Output press.} = (10 + a) + (10 + n) - (10)$$
$$= 10 + (a + n) = 10 + c \quad \ldots\ldots [1]$$

Therefore

$$c = a + n \quad \ldots\ldots [2]$$

The pneumatic controller was of the stacked-diaphragm type. Its output was connected directly to the process. This process consisted of three resistances and three small pneumatic tanks in series. Each tank was one tenth the volume of the preceding one. As a result, the process could be represented by a cascade of three single-time-constant noninteracting elements. An experimental frequency-response analysis made upon each tank individually and upon all three in cascade justified this simple representation.

Statham differential-pressure strain-gage transducers in conjunction with Brush analyzers were used to record n, c, and m. Water-pressure fluctuations were recorded on a circular chart by

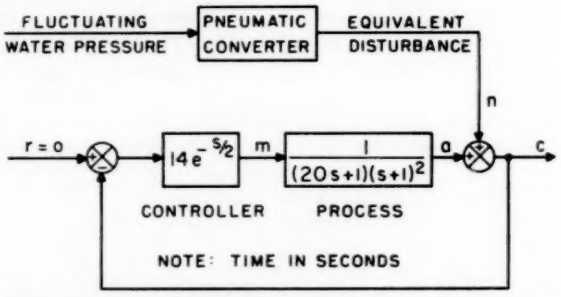


Fig. 2 Block Diagram of Pneumatic Pressure Control System

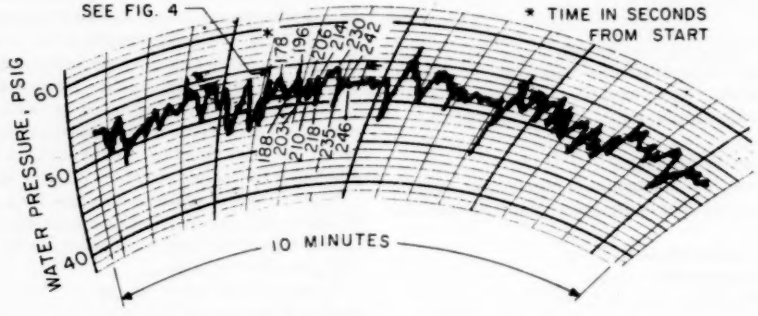


Fig. 3 Complete Water-Pressure Record; Keyed to Fig. 4

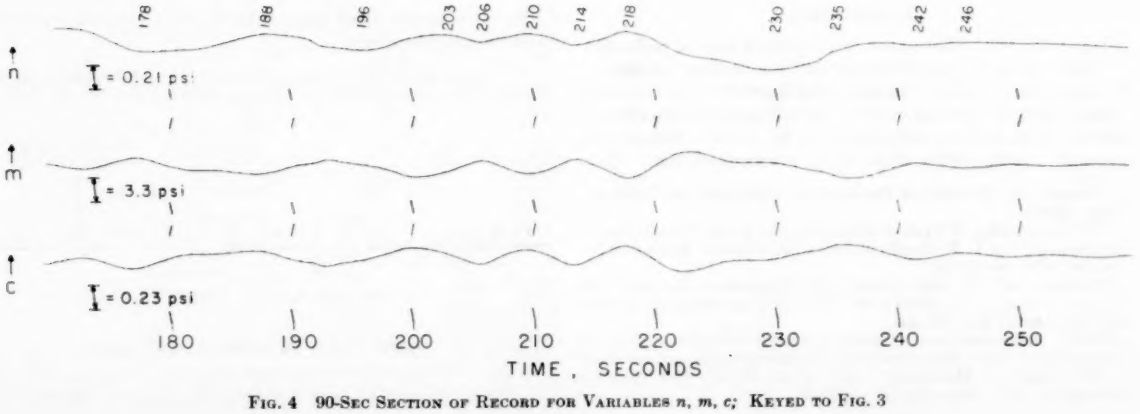


Fig. 4 90-Sec Section of Record for Variables n, m, c; Keyed to Fig. 3

a standard bellows-type pressure recorder. In all, 10 min of data were collected. The complete water-pressure record is reproduced in Fig. 3. A 90-sec section of the records for n, m, and c is shown in Fig. 4 with an indication of the way the one record keys into the other. Corresponding maxima and minima on the water-pressure record and the tape for n have been given the same numbers. Maxima on the water-pressure record correspond to minima on the record for n because the pneumatic converter was reverse-acting.

Calculation of Correlation Curves for Mathematical Model

Purely Random Disturbance (White Noise). The autocorrelation function,[4] $\phi_{nn}(\tau)$, for this type of disturbance is an impulse. Chosen for convenience is the unit impulse, represented by curve (1), Fig. 5. The corresponding spectral density is given by

$$\Phi_{nn}(\omega) = 2\int_0^\infty \phi_{nn}(\tau)\cos\omega\tau d\tau = 1 \quad \ldots\ldots\ldots [3]$$

Equation [3] shows that the spectral-density curve for white noise is a horizontal straight line which means that the disturbance has all frequencies, all at the same power level. $\Phi_{nn}(\omega)$ is an even function; that is, $\Phi_{nn}(\omega) = \Phi_{nn}(-\omega)$.

Forcing and response functions n and m are related through the transfer function

$$\frac{-2(s+1)}{s^2+s+2} \quad \ldots\ldots\ldots [4]$$

The spectral density of m is therefore

$$\Phi_{mm}(\omega) = \left|\frac{2(j\omega+1)}{(j\omega)^2+j\omega+2}\right|^2 \Phi_{nn}(\omega) = \left|\frac{2(j\omega+1)}{(j\omega)^2+j\omega+2}\right|^2 \quad \ldots\ldots [5]$$

$\Phi_{mm}(\omega)$, curve (3) of Fig. 5, is also an even function. Autocorrelation function, $\phi_{mm}(\tau)$, may now be found by evaluating integral Equation [6]

$$\phi_{mm}(\tau) = \frac{1}{\pi}\int_0^\infty \Phi_{mm}(\omega)\cos\tau\omega d\omega \quad \ldots\ldots\ldots [6]$$

In this evaluation, $\Phi_{mm}(\omega)$ was approximated by eight straight-line segments between $\omega = 0$ and $\omega = 8$ radians/unit time. Integration yielded

$$\phi_{mm}(\tau) = \frac{1}{\pi\tau^2}[0.625 + 2.50\cos 0.4\tau + 4.48\cos 0.8\tau - 7.60\cos 1.26\tau - 6.25\cos 1.44\tau + 3.75\cos 2\tau + 1.965\cos 2.6\tau + 0.4725\cos 4\tau + 0.0625\cos 8\tau] \quad \ldots\ldots [7]$$

Terminals of the straight-line-segment approximation of $\Phi_{mm}(\omega)$ are indicated by the coefficients for τ in Equation [7]. $\phi_{mm}(\tau)$ is again an even function and is shown by curve (5) of Fig. 5.[5]

[4] The autocorrelation function of a variable n is defined

$$\phi_{nn}(\tau) = \lim_{L\to\infty}\frac{1}{2L}\int_{-L}^{L} n(t)n(t+\tau)dt$$

[5] The block-diagram inversion procedure described in the section, Method for Discovering Disturbance Patterns, could have been used in this instance to calculate $\phi_{mm}(\tau)$. Instead, the route via the spectral-density calculation was chosen (*a*) to illustrate this particular method and (*b*) to show the correlation between the location of the spectral-density peaks of $\Phi_{mm}(\omega)$ and the minima of $\phi_{mm}(\tau)$, as discussed in the section, Analysis of Correlation Curves. By the same token, the spectral-density-type calculation could be substituted for the block-diagram inversion procedure if the integrals involved converge within a reasonable distance.

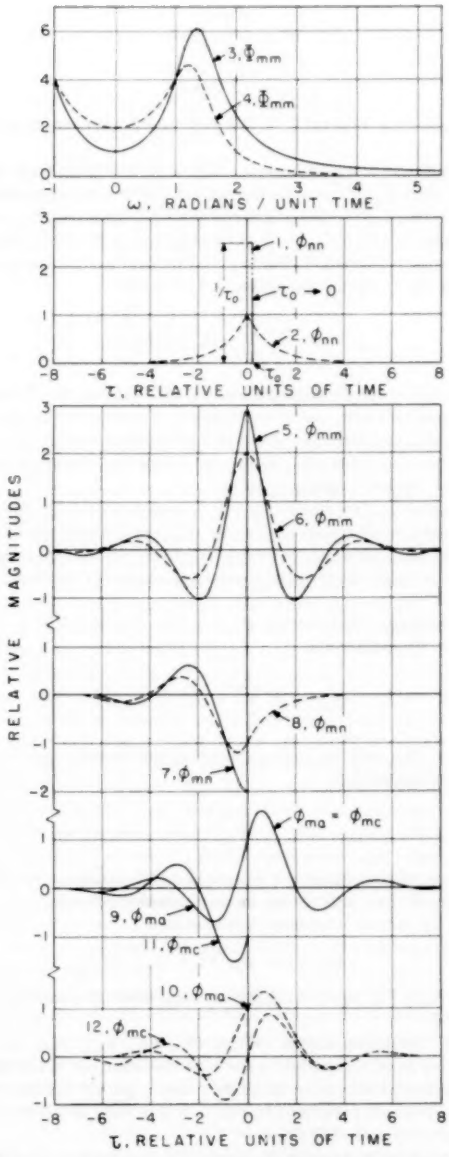


FIG. 5 SPECTRAL DENSITY AND CORRELATION CURVES FOR MATHEMATICAL MODEL

Next, crosscorrelation function[6] $\phi_{nm}(\tau)$ was evaluated by using the approximate expression for the convolution integral

$$\phi_{nm}(\tau) \simeq \sum_{k=0}^{p} g_k\phi_{nn}(\tau - kT) \quad \ldots\ldots\ldots [8]$$

Here $g(t)$ is the unit impulse response of the system described by Transfer Function [4] and

[6] The crosscorrelation function between two variables, n and m, is defined

$$\phi_{nm}(\tau) = \lim_{L\to\infty}\frac{1}{2L}\int_{-L}^{L} n(t)m(t+\tau)dt$$

$$g_0 = \left(\frac{T}{2}\right) g(0) \quad [9]$$

$$g_k = (T)g(kT) \quad [10]$$

A relative time interval $T = 0.2$ was used in all of these calculations.

Since $\phi_{nn}(\tau)$ is a unit impulse, at $\tau = 0$, it follows from Equation [8] that $\phi_{nm}(\tau)$ must be the same as the unit impulse response of the system, $g(t)$. This observation leads to a fruitful conclusion; namely, that $\phi_{nm}(\tau)$ may be regarded as the time response of the system subjected to a forcing function in time corresponding to $\phi_{nn}(\tau)$. When Equation [11] is written

$$m(t) \simeq \sum_{k=0}^{p} g_k n(t - kT) \quad [11]$$

the similarity between Equations [8] and [11] confirms the correctness of this very useful conclusion. A greater appreciation of these equations may be obtained by referring to the paper[3] by Goodman and Reswick. They also stress the fundamental importance of the foregoing concept.

A plot of $\phi_{mn}(\tau)$, curve (7), Fig. 5, was obtained by using the convenient relationship $\phi_{nm}(\tau) = \phi_{mn}(-\tau)$ which implies that ϕ_{nm} and ϕ_{mn}, although not even functions, are mirror images of each other with the reflection occurring about the vertical axis at $\tau = 0$.

Crosscorrelation function $\phi_{ma}(\tau)$ may be found in a similar manner, by evaluating

$$\phi_{ma}(\tau) \simeq \sum_{i=0}^{q} h_i \phi_{mm}(\tau - iT) \quad [12]$$

with $h(t)$ the unit impulse response of the process, described by the transfer function

$$\frac{1}{s+1} \quad [13]$$

However, the simplicity of Expression [13] makes it possible to generate $\phi_{ma}(\tau)$ directly by an easy graphic process.

Finally, since $c = a + n$, it follows that

$$\phi_{mc}(\tau) = \phi_{ma}(\tau) + \phi_{mn}(\tau) \quad [14]$$

Curve (11), Fig. 5, shows $\phi_{mc}(\tau)$ as computed by means of Equation [14].

Four correlation curves of interest, ϕ_{mm}, ϕ_{ma}, ϕ_{mn}, and ϕ_{mc}, have now been computed to describe the behavior of the mathematical model subjected to white noise, a purely random variation. These are portrayed in Fig. 5, where they are identified by odd numbers and solid lines.

Disturbance Having an Exponentially Decaying Autocorrelation Function. Equation [15] describes a disturbance having an exponentially decaying autocorrelation function

$$\phi_{nn}(\tau) = e^{-|\tau|} \quad [15]$$

This function, shown graphically by curve (2) of Fig. 5, decays in the same length of time as $h(t)$, the unit impulse response of the process.

A new set of correlation functions, ϕ_{mm}, ϕ_{ma}, ϕ_{mn}, and ϕ_{mc} was computed using the methods described in the foregoing. The resultant curves appear in Fig. 5 where they are identified by even numbers and dashed lines.

The spectral density for the equivalent disturbance is now

$$\Phi_{nn}(\omega) = 2\int_0^\infty e^{-|\tau|} \cos \omega\tau d\tau = \frac{2}{\omega^2 + 1} \quad [16]$$

which, if the values were plotted, would have the appearance of a bell-shaped curve. This means that there is no power at very high frequencies. The spectral density for m is

$$\Phi_{mm}(\omega) = \left|\frac{2(j\omega + 1)}{(j\omega)^2 + j\omega + 2}\right|^2 \left|\frac{2}{\omega^2 + 1}\right| \quad [17]$$

and is represented by curve (4) of Fig. 5.

Calculation of Correlation Curves for Real Process

The 10-min records of m, c, and n were read every 2 sec, yielding 300 ordinates for each variable. These data were then processed using a digital computer which solved discrete forms of the correlation functions such as, for example

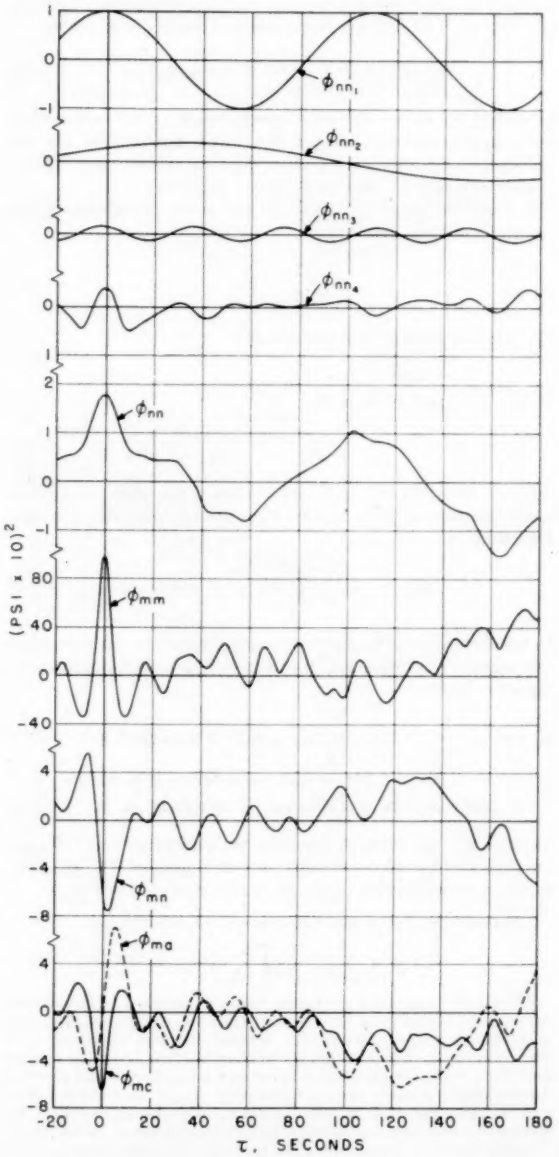


Fig. 6 Correlation Curves for Pressure Control System

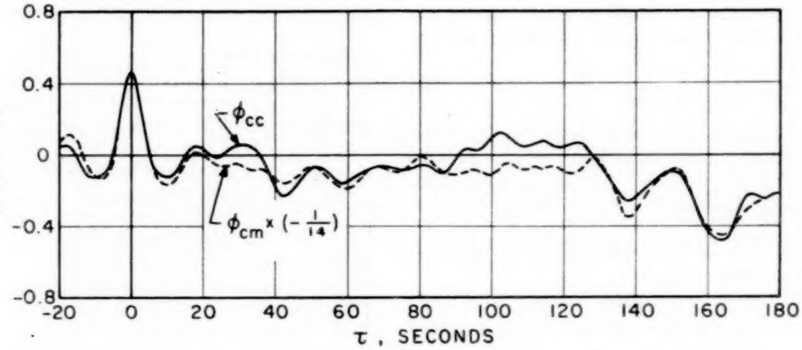


Fig. 7 Correlation Curves for Finding Pneumatic-Controller Transfer Function

$$\phi_{nn}(10) = \frac{1}{295}\,[n_1n_6 + n_2n_7 + n_3n_8 + \ldots\ldots + n_{295}n_{300}] \quad [18]$$

$$\phi_{mc}(4) = \frac{1}{298}\,[m_1c_3 + m_2c_4 + m_3c_5 + \ldots\ldots + m_{298}c_{300}] \quad [19]$$

A maximum value of τ equal to 180 sec or to 30 per cent of the total record length was chosen. This selection produced 90 calculated points at a spacing of 2 sec for each correlation curve.

Initially, calculated ordinates of correlation functions were relative to the edges of the record tapes involved. Also calculated by the digital computer was the average ordinate for each tape, allowing computation of the average or d-c component of all correlation functions. Subtracting this quantity from correlation function ordinates referenced to the edge of the tape left the desired correlation-function component. Figs. 6 and 7 show the correlograms ϕ_{nn}, ϕ_{mm}, ϕ_{mn}, ϕ_{mc}, ϕ_{cc}, and ϕ_{cm} obtained in this manner. Correlogram ϕ_{ma}, drawn dashed in Fig 6, was found by subtracting ϕ_{mn} from ϕ_{mc}.

Correlogram ϕ_{nn}, between $\tau = -20$ sec and $\tau = 180$ sec, was decomposed by trial and error into periodic components plus a remainder. No attempt was made to produce even functions only. The decomposition served to focus attention on the strong low-frequency periodic component, ϕ_{nn_1}, and on the peaks of the remainder, ϕ_{nn_4}, at $\tau = 0$ and at $\tau = 175$ sec.

Analysis of Correlation Curves

It is from an analysis of the correlation curves obtained for the mathematical model and those calculated for the real process that a knowledge of the effects of the disturbance pattern upon the system response emerges. As Goodman and Reswick[2] point out, with a purely random disturbance the dynamic relationship between ϕ_{mm} regarded as input and ϕ_{mc} regarded as output, for positive values of τ only, specifies the looked-for transfer function of the process. This is apparent from an inspection of curves (9) and (11) of Fig. 5, which show that ϕ_{mc} merges with ϕ_{ma} for positive values of τ. In other words, ϕ_{mc} is here identical to ϕ_{ma}, so that ϕ_{mc} may be regarded as the actual response of the dynamic-element coupling m and a, with ϕ_{mm} considered as the forcing function.

Also, Goodman and Reswick caution that the effect of the disturbance can extend a short distance into the positive τ region of ϕ_{mc}. Reference to curves (8), (10), and (12) of Fig. 5, and Equation [14], show this to be the case. However, the curves for ϕ_{mn}, ϕ_{ma}, and ϕ_{mc} in Fig. 6 for the real process cast some doubt upon their implication that this distance may usually be approximated.

Figs. 5 and 6 show that as the disturbance correlograms become wider and lower relative to the dynamics of the process block and hence depart from the shape of a single impulse at the origin, their effect persists for a greater distance to the right of $\tau = 0$. If an additional peak exists far from the origin, such as, for example, the peak at $\tau = 175$ sec in the remainder curve of ϕ_{nn}, namely, ϕ_{nn4} in Fig. 6, then ϕ_{mn} is again deviated from zero. This prolongs the separation of ϕ_{ma} and ϕ_{mc}. Since the region over which ϕ_{ma} and ϕ_{mc} coincide cannot be approximated without knowing the disturbance pattern to be expected, methods of ascertaining the latter become very important indeed.

But before turning to this consideration, Fig. 7 should be examined as an example in which the disturbance pattern does not intrude. Correlograms ϕ_{cc} and ϕ_{cm} may be thought of as the input and output, respectively, of the pneumatic controller. The curve for ϕ_{cm} in Fig. 7 is shown inverted and reduced by a factor of 14, the controller gain, to facilitate comparison with ϕ_{cc}. The two curves coincide or are close together for the major part of their course. Some discrepancy does exist. The cause for this has not been established as yet. A half-second dead time discovered in the controller by means of a frequency-response test is, of course, beyond resolution by these curves. Thus, the controller transfer function, as deduced from the two curves, would be -14; that is, $m = -14c$. This relationship is correct if we ignore the half-second dead time.

In addition, it should be noted that the correlation curves in Figs. 5 and 6 may be used to estimate the period of oscillation and the degree of damping of the control system. For example, the distance between the minima of ϕ_{mm}, curve (5), Fig. 5, is 4.0 relative units of time. For curve (6) it is 4.7. Significantly, the period of the damped oscillation of the mathematical model is also 4.7 relative units of time. The distance between the corresponding minima of the ϕ_{mm} curve for the real process is 17 sec, and a 16–17-sec period, the oscillatory period of the real process, is discernible in the record of m in Fig. 4. It is noted as a matter of interest that the maxima for the spectral-density curves of m, Fig. 5, occur at a circular frequency corresponding to the period of oscillation evident in the associated ϕ_{mm} correlograms.

Another and related observation from the same curves concerns the degree of damping in the system. The ratio between the ordinates of the central maximum of the ϕ_{mm} curve and the ordinate of the next minimum allows an estimate of the logarithmic decrement. This, in combination with a knowledge of the approximate period of oscillation of the system, enables an approximate damping ratio to be calculated. Application of this method to curve (6) of Fig. 5 produces the equation

$$\frac{0.6}{2.0} = e^{-\alpha\left(\frac{4.7}{2}\right)} \quad [20]$$

where α is the rate at which the oscillation is attenuated. Solution of Equation [20] yields the value of α as 0.5 per unit time. Using the relation

$$\alpha = \zeta\omega_{un} \quad [21]$$

where ζ = damping ratio, and ω_{un} = undamped natural circular frequency in radians/unit time, Equation [21] may be solved approximately for the damping ratio by setting $2\pi/4.7$ as a rough value for ω_{un}, thereby obtaining $\zeta = 0.37$. The actual value is $\zeta = 0.35$.

Testing this method on the real process, the following equation results from the ϕ_{mm} curve of Fig. 6

$$\frac{35}{98} = e^{-\alpha\left(\frac{17}{2}\right)} \quad [22]$$

The approximate value of α thus found is 0.12 sec^{-1}, and the corresponding approximate value of the damping ratio is $\zeta = 0.33$ No actual value of ζ was computed for the real process although its value was probably about 0.4.

Finally, the effectiveness, in a mean-square-error sense, with which the real control system was able to combat the injected disturbances is indicated by the central peaks of the ϕ_{nn} curve of Fig. 6 and the ϕ_{cc} curve of Fig. 7. The square root of the peak values, at $\tau = 0$, of autocorrelation curves ϕ_{nn} and ϕ_{cc} yields

$$\text{rms value of } n = 0.1\sqrt{1.8} = 0.13 \text{ psi}$$

$$\text{rms value of } c = 0.1\sqrt{0.45} = 0.067 \text{ psi}$$

Thus, the control system cut in half the rms value of the potential pressure deviation from the control point caused by the disturbance.

Method for Discovering Disturbance Patterns

Discovery of the disturbance pattern is straightforward, at least theoretically, if equivalent linear dynamic characteristics are available for all elements in the control loop, and ϕ_{mm} has been calculated from a recording of m. For example, the dynamic relationship between input n and response m for the pneumatic process is given by the transfer function of Fig. 8. Inversion of the transfer function of Fig. 8 yields the transfer function of Fig. 9 which specifies the dynamic relationship between m, now regarded as input, and n, now considered to be output rather than input. From this it follows that ϕ_{mm}, regarded as input to the dynamic element specified by Fig. 9 will yield ϕ_{mn} as output.

It should be noted here that since ϕ_{mn} represents the difference between ϕ_{ms} and ϕ_{mc}, the ϕ_{ms} correlogram is already useful for estimating the distance into the positive τ-region of ϕ_{mc} that the disturbance effect cannot be ignored. Calculations to this point, based upon industrial records of m extracted from process-control systems with known dynamic characteristics, should be of considerable aid to control engineers, provided results of many different tests are published.

Proceeding with the method for discovering ϕ_{nn}, Fig. 8 specifies the dynamic relationship between ϕ_{nn} regarded as input and ϕ_{nm} regarded as output. With ϕ_{mn} known, simple inversion about the vertical or $\tau = 0$ axis yields ϕ_{nm}. Thus the output of the system of Fig. 8 is known. Inverting the transfer function of Fig. 8 to yield the transfer function of Fig. 9 allows the original output ϕ_{nm} to be treated as input to the dynamic element specified by Fig. 9. The resulting output is the sought-for autocorrelation function of the disturbance; namely ϕ_{nn}.

In short, the steps for discovering ϕ_{nn} may be listed as follows:

1 Calculate ϕ_{mm} from a record of m.

2 Derive the dynamic relationship between n, regarded as

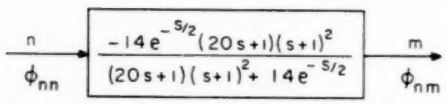


Fig. 8 Transfer Function Relating n and m, also ϕ_{nn} and ϕ_{nm}

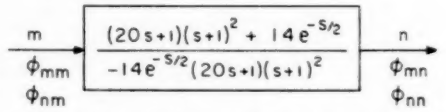


Fig. 9 Inverse Transfer Function Relating m and n, ϕ_{mm} and ϕ_{mn}, and ϕ_{nm} and ϕ_{nn}

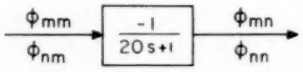


Fig. 10 Approximate Inverse Transfer Function

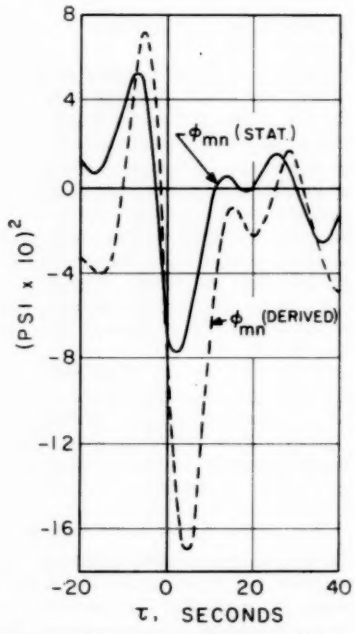


Fig. 11 Statistically Calculated ϕ_{nn} and ϕ_{nn} Derived From ϕ_{mm}

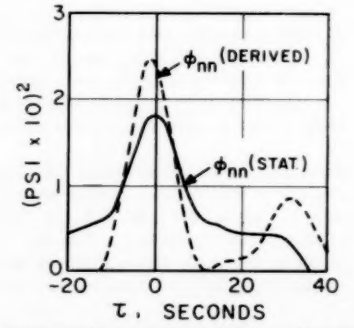


Fig. 12 Statistically Calculated ϕ_{nn} and ϕ_{nn} Derived From ϕ_{nm}

input, and m, regarded as output, from a knowledge of the control system.

3 Define a new dynamic element with characteristics the inverse of those found in step (2).

4 Force ϕ_{mm} through this element, either by mathematical calculation, or physically, using an analog, and get the response ϕ_{mn}.

5 Rotate ϕ_{mn} about the vertical $\tau = 0$ axis to produce ϕ_{nm}.

6 Force ϕ_{nm} through the same dynamic element specified in step (3) and obtain finally the desired function ϕ_{nn}.

Laning and Battin[7] describe a similar procedure in their recently published book to which the reader is referred for a wealth of information on the statistical approach to the control problem.

If analog techniques are to be used for effecting the inversion of a dynamic element, at least two methods are available. One method simulates the inverse transfer function directly. The other simulates the direct transfer function and uses it in the feedback path around a high-gain amplifier, thereby producing an over-all transfer function with the desired inverse characteristics. Simulation of the direct-transfer function may be accomplished either by setting the computer to perform the desired operations relating n and m, or by simulating separately each element in the control loop before linking them.

A rough demonstration of the foregoing procedure may be made if the system of Fig. 9 is approximated by the simple system shown in Fig. 10. With ϕ_{mm} as input to the latter a graphic solution yielded the dashed curve ϕ_{mn} of Fig. 11. For comparison the original ϕ_{mn} curve, drawn solid, is also shown.

Repeating the graphic process with ϕ_{nn} as input (obtained by reversing ϕ_{mn} of Fig. 6 about the $\tau = 0$ axis) the dashed line in Fig. 12 was found. To avoid accumulating errors the graphically derived ϕ_{nn} curve of Fig. 11 was not used. The statistically calculated ϕ_{nn} curve is also shown in Fig. 12.

At this point, it should be stressed that ϕ_{nn} represents the autocorrelation function of an equivalent disturbance which may be regarded as the sum of all the individual disturbance effects upon controlled variable c produced by several different disturbances entering the control loop between m and c. It is this equivalent disturbance for which the controller attempts to correct.

The primary function of process controllers is to minimize the effect of disturbances upon the controlled variable. Usually the controller settings are arrived at from stability considerations of the control loop. Controller settings based, instead, upon a consideration of the equivalent disturbance pattern to be expected may result in better performance. The difficulty to date has been the discovery and description of the disturbance pattern. It is here suggested that the discovery of the autocorrelation function of the equivalent disturbance in the manner indicated may help to surmount the stated difficulty. Should this method for discovering equivalent disturbance patterns prove successful, better controller settings should be possible than are now used. Also, a knowledge of equivalent disturbance patterns will allow better judgment in the use of correlation functions obtained from normal operating records for dynamic analysis.

Conclusions

1 Correlation functions obtained from normal operating records provide a powerful tool for dynamic analysis of control systems without introducing additional disturbances.

2 These correlation functions, however, should be used with caution when disturbances enter the control loop between the two points defining input and output of the unknown element to be analyzed.

3 Knowledge of the equivalent disturbance patterns is important (*a*) for successful use of correlation functions obtained from normal operating records and (*b*) for obtaining better controller settings.

4 Block-diagram inversion is suggested as a method for discovering the equivalent disturbance pattern entering a control loop.

5 Despite the value of correlation techniques for analysis when additional disturbances cannot or should not be introduced, frequency-response data obtained under favorable conditions, and when disturbances can be introduced, will yield more accurate transfer functions.

Acknowledgments

Research upon which this paper is based was supported by a grant from the Institute of Engineering Research at the University of California, Berkeley, Calif. The author wishes to express his appreciation to Mr. J. M. Maughmer and other staff members of the Digital Computer Laboratory, Cory Hall, U. C. Campus. Thanks are due also to three students: Messrs. L. R. A. Austin, and R. D. Davis, and especially to Mrs. Nancy Dabaghian who assisted in the many calculations underlying this paper. It was the enthusiastic interest of Yasundo Takahashi, Professor of Mechanical Engineering at the University of Tokyo, in optimum control settings, which led the author to explore the possibilities of finding noise patterns. This exploration, in turn, developed into a consideration of the effect of these patterns upon the findings of Prof. T. P. Goodman and Prof. J. B. Reswick of the Massachusetts Institute of Technology. It was from their studies of statistical techniques for analyzing process-transfer functions that the author, demonstrably, derived many fruitful clues, indeed the basic foundation of his thinking on this subject.

Discussion

Thomas P. Goodman.[8] This paper is a welcome addition to the growing body of literature on statistical methods in automatic-control problems. The equivalent disturbance patterns described by the author will be a useful tool for specifying more precisely the region of the cross-correlation curve (referred to in Appendix 1 of the paper by Reswick and the writer[2] as the region of $\tau > A$) to be used in the process of deconvolution to determine the dynamic characteristics of a system.

The author's Equations [18] and [19] for computing correlation functions are not quite the same as those given by Reswick and the writer.[2] Strictly speaking, to make Equations [8] and [12] valid, each point on the correlation curves should be based on the same number of ordinates. Since only 210 ordinates of m, n, and c can be used in computing the last point on each correlation curve, it can be argued that only 210 ordinates should be used in computing the other points as well. Using this argument, Equations [18] and [19] should be rewritten

$$\phi_{nn}(10) = \frac{1}{210}\left[n_1 n_6 + n_2 n_7 + n_3 n_8 + \ldots + n_{210} n_{215}\right]$$

$$\phi_{mc}(4) = \frac{1}{210}\left[m_1 c_2 + m_2 c_4 + m_3 c_5 + \ldots + m_{210} c_{212}\right]$$

However, the error introduced by using a greater number of ordinates in computing these points on the correlation curves is probably small, and the author's method has the advantage of utilizing more of the information available in the original records.

It is known that spurious effects can arise in correlation curves computed on the basis of only a few hundred ordinates, and these

[7] "Random Processes in Automatic Control," by J. H. Laning, Jr., and R. H. Battin, McGraw-Hill Book Company, Inc., New York, N. Y., 1956, p. 219.

[8] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Assoc. Mem. ASME.

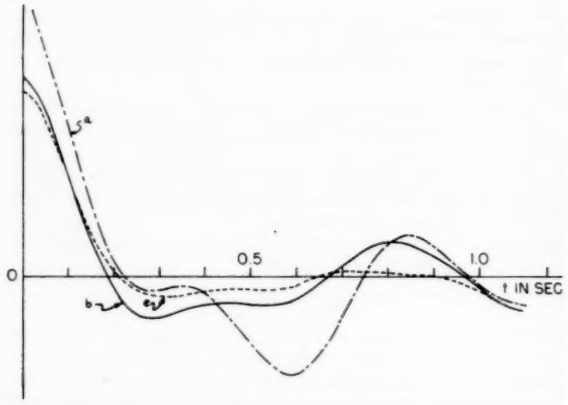


FIG. 13 COMPUTED AUTOCORRELATION FUNCTIONS FOR DIFFERENT LENGTHS OF SAME RECORD

(Points taken at 0.01-sec intervals.)

(*a*, 4-sec record; *b*, 12-sec record; *c*, 21-sec record.)

effects may account for some of the oscillations in the correlation functions of Fig. 6 for $\tau > 20$ sec, and for the discrepancies in the region $\tau > 20$ sec between the solid and dotted curves at the bottom of Fig. 6 and in Fig. 7. To illustrate these spurious effects, Fig. 13 shows autocorrelation curves computed on the basis of three different numbers of ordinates from the same original record.[9] If the curve based on the largest number of ordinates is assumed to be closest to the true autocorrelation function (as defined in footnote 4 of the paper), it can be seen that the curves computed on the basis of smaller numbers of ordinates contain spurious oscillations. It would be highly desirable to find a way of computing modified correlation functions in which these spurious oscillations would be damped out.

The writer would like to see a more detailed explanation of the author's method for determining the period of oscillation and degree of damping of the system from the correlation curves. It would be interesting to know how generally applicable this method is.

OTTO J. M. SMITH.[10] Mr. Thal-Larsen is to be commended upon the excellent treatment of this subject, and his confirmation of the relationship between the autocorrelation functions measured in a closed loop, and the transferences of the loop. Mr. Thal-Larsen has demonstrated that a measurement of the autocorrelation and the cross correlation of the output of a process, and the output of the controller on a continuously operating process can be used to yield the transference of the process, if the equivalent disturbance at the output is a flat gaussian random spectrum. Since actual disturbances may enter within the process, the equivalent disturbance at the output is more likely to be a filtered random gaussian spectrum, with an autocorrelation function more like curve 2 in Fig. 5.

The Two-Test Method

If the spectrum of the equivalent disturbance is not known, and if the process transference is not known, then measurements of the various autocorrelations and cross correlations which are available in the closed loop are not sufficient for one to be able to solve for the process transference. It is necessary for measurements to be made for two different operating conditions: One may measure the correlation functions for two different controller settings, or one may measure the correlation functions for two different signal and disturbance spectra. In both cases, the operation of the loop is altered. In the first case, the loop gain can be reduced, in which case the deviation of the output from the reference is increased. In the second case, one may introduce a flat random gaussian signal at the reference point and measure the effect of this signal throughout the system.

Power-Density Spectra

The measurements which must be made on the actual process are recording the output variable and the input variable of the process for a long length of time on a magnetic tape or other recording medium. These recorded signals must then be processed in the laboratory, by playing them through an autocorrelation-function computer which measures the average product of the signal times the signal displaced by some fixed quantity of time. These autocorrelation functions can be plotted much as the author has done in Fig. 5. The autocorrelation function is a function of time and therefore difficult to handle mathematically when one is solving for the transference of a process control loop, whose Laplace transform is a function of frequency. Therefore this discusser prefers to convert the measured autocorrelation functions directly into power-density spectra by taking their Fourier transforms

$$\phi_{11}(\lambda) = \mathfrak{F}\psi_{11}(\tau) = \int_{\tau=-\infty}^{+\infty} d\tau \epsilon^{-j\lambda\tau}\psi_{11}(\tau) \ldots\ldots [23]$$

$$\psi_{11}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} f_1(t)f_1(\tau + t)dt \ldots\ldots [24]$$

λ is a complex frequency variable usually denoted by

$$\lambda = \omega + ja \ldots\ldots [25]$$

$\phi_{11}(\lambda)$ is the power density spectrum, which is a ratio of polynomials in λ. It has units of signal squared per cycle per second. The numerator polynomial can be factored into its roots and these roots are called the zeros of the spectrum. The denominator polynomial can be factored into its roots and these are called the poles of the spectrum. Both the zeros and the poles of a self-power spectrum are located on the corners of rectangles in the complex λ plane.

The Fourier transform of the cross correlation between one signal and another is called the cross-power density spectrum between the two signals. The spectral input to a transference times the transference expressed as the Fourier transform of its weighting function (impulse response) is equal to the cross-power density spectrum between the input and the output of the transference. The output-power density spectrum of any operational device is equal to the input-power density spectrum of that operational device times the transference and times the complex conjugate of the transference. With these rules one can write down immediately the relationship between the spectrum at one point in a system and the spectrum at a different point in the same system.

[9] These curves were obtained by C. M. Chang at the writer's suggestion, using an electronic digital computer. The recorded variable was a random voltage obtained by passing the output of a random noise generator through a low-pass filter. See "A New Technique of Determining System Characteristics From Normal Random Operating Records," by C. M. Chang, Mechanical Engineer's Thesis, Massachusetts Institute of Technology, January, 1955; "Experimental Determination of System Characteristics From Correlation Measurements," by T. P. Goodman, ScD Thesis, Massachusetts Institute of Technology, June, 1955. See also "Contributions to the Study of Oscillatory Time-Series," by M. G. Kendall, Cambridge University Press, Cambridge, England, 1946, chapter 3; "An Introduction to Stochastic Processes," by M. S. Bartlett, Cambridge University Press, 1955, chapter 9.

[10] Professor of Electrical Engineering, University of California, Berkeley, Calif.

Variable Gain Method

With reference to the author's diagram of the control system in Fig. 1, I would like to use the notation that the transference of the controller from the error to m is G and that the transference of the process from m to a is H. The reference shall be kept constant so that all of the statistical fluctuations in the output shall be due to the equivalent disturbance. The first test shall be denoted by subscripts 1 and shall be made with the highest possible loop gain for which the system is stable and operates relatively satisfactorily. The second test shall be denoted by subscripts 2 and shall be performed with a loop gain several times less than the value used for test 1. The output self-power spectrum for the first test is

$$\phi_{cc1}(\lambda) = \frac{1}{(1 + G_1H)\overline{(1 + G_1H)}}\,\phi_{nn}(\lambda) \quad [26]$$

The output self-power spectrum for test 2 is

$$\phi_{cc2}(\lambda) = \frac{1}{(1 + G_2H)\overline{(1 + G_2H)}}\,\phi_{nn}(\lambda) \quad [27]$$

In these equations, the bar over the transference function means the complex conjugate of the phasor transference expressed as a function of the complex frequency variable $\lambda = \omega + ja$. Each of these self-power spectra or functions can be expressed as a ratio of polynomials in λ and these polynomials can each be factored into upper λ-half-plane roots and lower λ-half-plane roots. The upper λ-half-plane roots contribute to the positive τ portions of the autocorrelation function, and the lower λ-half-plane roots contribute to the negative τ portions of the autocorrelation function. Since both Equations [26] and [27] are completely symmetrical with no phase, and every upper λ-half-plane pole is matched by its mirror image in the lower λ-half plane, it is possible to factor Equations [27] and [26] into their upper λ-half-plane functions times the lower λ-half-plane functions. These are represented respectively by $+$ and $-$ superscripts referring to positive-time poles and negative-time poles. This factorization for Equation [26] is shown in Equations [28] and [29]

$$\phi_{cc1}(\lambda) = \phi_{cc1}{}^{+}(\lambda)\cdot\phi_{cc1}{}^{-}(\lambda) \quad [28]$$

$$\phi_{cc1}(\lambda) = \frac{\phi_{nn}{}^{+}(\lambda)}{(1 + G_1H)} \cdot \frac{\phi_{nn}{}^{-}(\lambda)}{\overline{(1 + G_1H)}} \quad [29]$$

Taking the upper-half λ-plane poles only for the first test one has

$$\phi_{cc1}{}^{+}(\lambda) = \frac{\phi_{nn}{}^{+}(\lambda)}{(1 + G_1H)} \quad [30]$$

taking the upper-half λ-plane poles only for the second test one has

$$\phi_{cc2}{}^{+}(\lambda) = \frac{\phi_{nn}{}^{+}(\lambda)}{(1 + G_2H)} \quad [31]$$

The ratio of these is

$$\frac{\phi_{cc1}{}^{+}(\lambda)}{\phi_{cc2}{}^{+}(\lambda)} = \frac{1 + G_2H}{1 + G_1H} \quad [32]$$

This can be solved for H and yields

$$H = \frac{\phi_{cc2}{}^{+} - \phi_{cc1}{}^{+}}{G_1\phi_{cc1}{}^{+} - G_2\phi_{cc2}{}^{+}} \quad [33]$$

$$G_1H = \frac{\dfrac{\phi_{cc2}{}^{+}}{\phi_{cc1}{}^{+}} - 1}{1 - \dfrac{G_2}{G_1}\left(\dfrac{\phi_{cc2}{}^{+}}{\phi_{cc1}{}^{+}}\right)} \quad [34]$$

This is the closed loop transference for the first test directly in terms of the ratio of the power density spectra at the output between the second test and the first test and the ratio of the transferences of the controller between the second test and the first test.

If the change in the controller is only a change of gain and does not involve a change in the location of the poles and zeros of the controller, then

$$G_2 = kG_1 \quad [35]$$

where k is a number much less than unity. For this special case the closed loop transference can be calculated by

$$G_1H(\lambda) = \frac{\left(\dfrac{\phi_{cc2}{}^{+}}{\phi_{cc1}{}^{+}}\right) - 1}{1 - k\left(\dfrac{\phi_{cc2}{}^{+}}{\phi_{cc1}{}^{+}}\right)} \quad [36]$$

The solution of the numerator and denominator roots for this equation can be carried out rapidly as a root locus plot on an s-plane analog. Either a conduction sheet electrical analog can be used, or the geometric analog of Walter Evans, utilizing a Spirule, may be used for this calculation.

The closed loop poles of the system are also available directly from the measured data. They are the roots of $1 + G_1H$ where

$$(1 + G_1H) = \frac{(1 - k)\phi_{cc2}{}^{+}}{\phi_{cc1}{}^{+} - k\phi_{cc2}{}^{+}} \quad [37]$$

The spectrum of the noise may also be calculated directly and the equivalent load disturbance at the output is given by

$$\phi_{nn}{}^{+}(\lambda) = \frac{1 - k}{\dfrac{1}{\phi_{cc2}{}^{+}} - \dfrac{k}{\phi_{cc1}{}^{+}}} \quad [38]$$

$$\phi_{an}(\lambda) = \phi_{nn}{}^{+}(\lambda)\cdot\overline{\phi_{nn}{}^{+}(\lambda)} \quad [39]$$

Variable Signal Method

The transference of a closed loop controller operating with an unknown statistical disturbance can be calculated directly from two tests, one in the absence of any extra input signal, and the second including the effect of a deliberately introduced input signal. The spectrum of the deliberately introduced signal at the reference point r in Fig. 1 must be known or measured. One must measure the self-power spectrum or the autocorrelation function of the input in Fig. 1 and of the output of the system for the two cases of the undisturbed system and the disturbed system. The first case when there is no input, yields an output-power spectrum given by

$$\phi_{cc1}(\lambda) = \frac{1}{(1 + GH)\overline{(1 + GH)}}\,\phi_{nn}(\lambda) \qquad r = i = 0 \quad [40]$$

For the second test when there is a deliberately introduced input, the output-power spectrum is given by

$$\phi_{cc2}(\lambda) = \frac{1}{(1 + GH)\overline{(1 + GH)}}\,\phi_{nn}(\lambda) + \left(\frac{GH}{1 + GH}\right)\overline{\left(\frac{GH}{1 + GH}\right)}\,\phi_{ii}(\lambda) \qquad r = i \neq 0 \quad [41]$$

The effect of the unknown equivalent disturbance at the output can be eliminated by solving for ϕ_{nn} in one equation and substituting it in the other. Eliminating ϕ_{nn}

$$[\phi_{cc2} - \phi_{cc1}] = \left(\frac{GH}{1 + GH}\right)\overline{\left(\frac{GH}{1 + GH}\right)}\,\phi_{ii} \quad [42]$$

Factoring these ratios of polynomials in λ into the upper λ-half-plane roots and the lower λ-half-plane roots, one has for the upper λ-half-plane roots

$$[\phi_{cc2} - \phi_{cc1}]^+ = \left(\frac{GH}{1 + GH}\right) \cdot \phi_{ii}^+ \quad \ldots\ldots\ldots [43]$$

This can be solved directly for the reciprocal loop transference

$$\frac{1}{GH} = \frac{\phi_{ii}^+}{[\phi_{cc2} - \phi_{cc1}]^+} - 1 \quad \ldots\ldots\ldots\ldots [44]$$

The loop transference is

$$GH = \frac{[\phi_{cc2} - \phi_{cc1}]^+}{\phi_{ii}^+ - [\phi_{cc2} - \phi_{cc1}]^+} \quad \ldots\ldots\ldots\ldots [45]$$

The procedure to be used for calculating the loop transference from Equation [45] is as follows:

1. Measure the output-autocorrelation function in the absence of the input and take its Fourier transform to yield the output-power spectrum.
2. Measure the output autocorrelation function in the presence of a known input and take its Fourier transform to yield the second output-power spectrum, $\phi_{cc2}(\lambda)$.
3. Measure the input autocorrelation function for the second case and take its Fourier transform to yield the input-power spectrum $\phi_{ii}(\lambda)$.
4. Take the difference between the second output spectrum and the first output spectrum, and solve for the new zeros of the function. This requires a root locus plot in the λ plane.
5. Take only the upper λ-half-plane poles and zeros for the numerator of Equation [45].
6. Solve for the denominator of Equation [45] with a second λ-plane root locus plot.

The closed loop roots are directly available from the measured power spectra and they are

$$(1 + GH) = \frac{\phi_{ii}^+}{\phi_{ii}^+ - [\phi_{cc2} - \phi_{cc1}]^+} \quad \ldots\ldots\ldots [46]$$

The power spectra of the equivalent disturbance at the output can also be solved for directly by substituting back into Equation [40] and this yields

$$\phi_{nn}^+(\lambda) = \frac{\phi_{ii}^+ \phi_{cc1}^+}{\phi_{ii}^+ - [\phi_{cc2} - \phi_{cc1}]^+} \quad \ldots\ldots\ldots [47]$$

Summary

It has been shown that the power spectrum of the equivalent disturbance entering an operating closed-loop process control, and both the open-loop and the closed-loop transferences of the control can be obtained directly from autocorrelation measurements of the signals at the output of the process and at the input to the controller. Two sets of measurements must always be made and the difference between the measurements can either be a change in the process or a change in the signals in the process. It is important to note that the particular measurements which have been chosen for these transference determinations do not require any cross-correlation calculations. It is possible to use the signal m in Fig. 1 instead of the input function for one of these tests, but this requires more complex mathematics. This discusser has chosen those measurements which are the easiest to make and which have the simplest mathematical steps to obtain the closed-loop transference.

YASHUNDO TAKAHASHI.[11] This paper contains most valuable contribution to the understanding of the feedback control processes under stochastic inputs. Relations between actual noise patterns in the system, such as presented in the paper, would constitute the direct basis for future system design and performance optimization. The discusser has been working on a similar topic since he had an opportunity of working with the author. As the work is still under way, only one result will be shown here, leaving others to a later occasion of publication. The optimum setting of a proportional controller gain k, which gives minimum square error, for a process of reaction rate R and a dead time L, when white noise is introduced as a disturbance at the manipulated variable side of the process, as shown in Fig. 14, can be found from Fig. 15 as follows

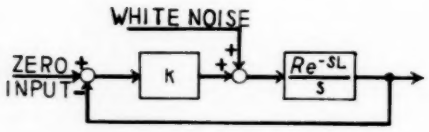


FIG. 14

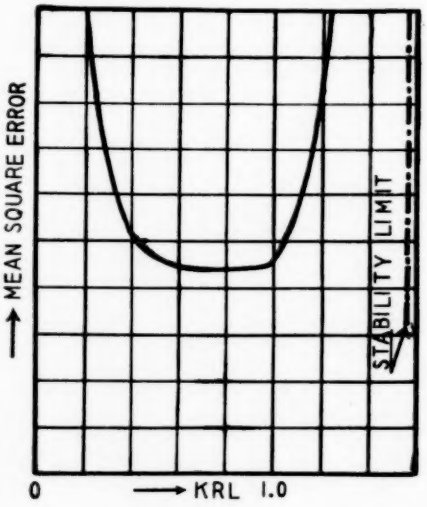


FIG. 15

$$kRL = 0.6 \text{ to } 1.0$$

As can be seen from Fig. 15, the mean-square error in this case is almost flat throughout the optimum range stated. It will be of interest to note that the maximum value of kRL just given coincides with Ziegler-Nichols' optimum setting.

The equivalent noise concept suggested in the original paper is very useful for handling stochastic problems of feedback control systems. However, the assumption of white noise as the equivalent noise pattern would lead to a difficulty in optimum setting investigations, due to the fact that the white noise extends indefinitely in high-frequency range. This is the reason why the writer assumed the system of Fig. 14 instead of using the concept of equivalent noise.

AUTHOR'S CLOSURE

The author appreciates the discussions prepared by Professors Goodman, Smith, and Takahashi. In his view, the value of this paper has been increased with their contributions.

Professor Goodman's discussion, mailed from Munich, Ger-

[11] Professor, Institute of Industrial Science, University of Tokyo, Chiba City, Japan.

many, was received by the author on the very eve of his own departure for a trip abroad. He thanks Professor Goodman for taking the time and trouble to prepare a discussion under difficult circumstances, and regrets the haste with which his own reply had to be prepared.

The author is well aware of the effect which the number of ordinates used for calculation of correlation functions can have on the shape of the latter, and he is happy to have it brought to the reader's attention. It is necessary to stress that the correlation curves for the real process are not statistically stationary. Thus curves computed from much longer records, perhaps hours or days in length, can be expected to be appreciably different from the ones shown. It is conceivable that, at the larger values of τ, the high-frequency oscillations will disappear and perhaps even the low-frequency components, as longer and longer records are used for calculating the correlation curves. Unfortunately, the figure presumably portraying this effect is not available at this writing, and its examination will have to be postponed until after publication.

The author agrees with Professor Goodman that the error introduced by not using the same number of ordinates in calculating the correlation curves of Figs. 6 and 7, is probably small. But he is puzzled by the suggestion that some of the oscillations in the correlation functions of Fig. 6 for $\tau > 20$ sec may be due to spurious effects. As the correlation functions in this paper are based upon a relatively large number of ordinates for the system in question, it does not appear likely that the oscillations are spurious. Since these correlation functions may be looked upon as time response curves, the continuing oscillations are accounted for by the natural oscillations of the control system which is continuously excited by the various high-frequency components present in the noise autocorrelation function. The low-frequency oscillation is the response of the system to the low-frequency components of the noise pattern.

Similar considerations led the author to regard the system response near $\tau = 0$, if produced by a relatively high and sharp central peak of the noise autocorrelation curve, as being predominantly the impulse response of the system. With this thought in mind he wrote Equations [20] and [22] as a rough approximation. The author agrees that it will be interesting to test this approximation on other correlation curves, especially if the noise pattern is also known.

Professor Smith points out the necessity for additional disturbance of the control system, over and above that originally present, to find transferences and the original disturbance pattern. The additional disturbance consists of either a change in the controller transference or of the injection of an additional signal at the controller. But, avoiding such additional disturbances was the very justification for the use of statistical techniques—techniques that require time-consuming data reduction methods if the calculations are not performed automatically by computers.

If we do admit that additional disturbances are needed to find transferences (and operating crews have been known to walk off the job in a huff leaving the plant entirely in the hands of the test engineers in instances where additional disturbances were introduced), the author believes there is a better way than that suggested by Professor Smith. In fact, the author believes the best method for finding transferences in the presence of noise that has been developed to date, is the one reported by Dr. P. E. A. Cowley[12] that combines advantages of both the frequency-response and the statistical techniques.

The reader interested in the two approaches outlined by Prof. Smith is referred to a paper by J. H. Westcott[13] in which he treats this subject in a similar vein. Both the Smith and Westcott approaches have merit in that the shape of the equivalent disturbance pattern does not affect determination of the process transference. At the same time, their two-test method has an inherent weakness in that it assumes the equivalent disturbance pattern will be the same (statistically stationary) for both tests. It is the author's experience that this is a dangerous assumption.

Professor Takahashi's preliminary report on optimum controller settings indicate a type of exploration which, although it has not as yet produced any startling results, may ultimately lead to better performance of control systems.

Takahashi's curve relating the mean square error to controller gain shows how steeply the former rises when the latter is changed appreciably from optimum. This type of evidence, plus the requirement that the plant product stay on specification, would argue against the usefulness of the variable-gain method proposed by both Smith and Westcott.

---

[12] Reference cited under footnote 3 and since published in Trans. ASME, vol. 79, 1957, pp. 823–832.

[13] "The Determination of Process Dynamics From Normal Disturbance Records of a Controlled Process'" by J. H. Westcott, Fachtagung, Regelungstechnik Heidelberg, Germany, 1956, Beitrag Nr. 40, Unkorrigierter Vorabdruck.

An Analog Study of a High-Speed Recording Servomechanism

By J. W. SCHWARTZENBERG,[1] PHILADELPHIA, PA.

This paper presents an analysis of a high-speed recording servomechanism. Because of the nonlinearities present, the system is simulated on an analog computer. Minimum balancing times with a 100 per cent step-input signal are obtained for various combinations of system parameters. Several types of damping are considered, and the a-c carrier action of the amplifier is included in the simulation. The frequency response and following error of the system are also considered.

Introduction

THE electronic recording servomechanism has become an important instrument for measuring many variables accurately. Although a wide variety of applications can be met with recorders having full-scale balancing times of from 1 to 5 sec, some of the faster processes in service today require full-scale balancing times of $^1/_2$ sec or less. One example of this is a multipoint scanning system for a high-speed wind tunnel where the recorder must come to balance within a prescribed time.

This paper presents an analysis of a high-speed recording servomechanism. It was undertaken to determine the minimum time of response for a 100 per cent step change in recorder input. One important performance criterion was that minimum response time must be obtained with a reasonable size motor and with practical values of load inertia and load friction.

Because of the nonlinearities present in this problem, such as amplifier limiting, dry-friction load, and the motor speed-torque curve, it was desirable to use an analog computer for the analysis. All results in this paper are obtained from the computer simulation with the exception of some actual test results presented at the end of the paper.

The importance placed upon transient response in most applications makes it necessary to choose this as the governing performance criterion. However, frequency response and following error also are considered.

Description of System

The system considered in this paper is a common form of recording servo in general use today. As shown in Fig. 1, the basic system consists of an amplifier, servomotor, gear train, feedback element, damping network, and input filter.

The amplifier is a high-gain d-c input, chopper-modulated, a-c output device which provides an a-c voltage for the control winding of the servomotor. Since the amplifier necessarily must have a finite output it can be seen that the voltage applied to the motor must be limited at some value. This limiting effect will be included in the analysis as well as the carrier effect of the amplifier.

The motor is a high-performance, low-inertia, 2-phase 60-cycle a-c servomotor. Although better results can be expected from a 400-cycle motor, it is felt that devices requiring 400-cycle power have limited industrial application at the present time. The motor constants of the Diehl 5-watt, 2-pole, FPE 25-22 servomotor are used for numerical values.

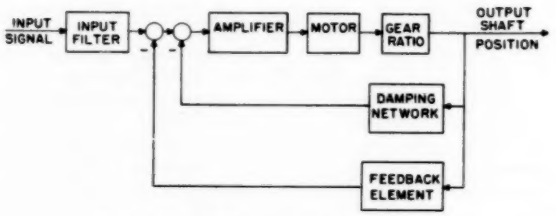


Fig. 1 Basic Block Diagram of System

The gear train is used to match the dry friction and the inertia characteristics of the load to the motor.

The feedback element converts the position of the output shaft into an electrical quantity to match the input signal.

The damping network is required to compensate for the lags present in the system. It is adjusted not only to give stable operation but to give the proper transient response. Several forms of damping are considered in this paper, lead-lag compensation for various values of gain ratio, linear velocity feedback, and absquare[2] velocity feedback.

The input filter is used to prevent large amounts of 60-cps pickup from entering the amplifier and causing insensitivity and overloading.

Computer Simulation

Amplifier. The saturation or limiting effect of the servo amplifier is assumed to be of the form shown in Fig. 2. It also can be described by the equations

$$\begin{aligned} e_2 &= Ke_1 && -E_{max} < Ke_1 < E_{max} \\ e_2 &= E_{max} && Ke_1 \geqslant E_{max} \\ e_2 &= -E_{max} && Ke_1 \leqslant -E_{max} \end{aligned} \quad \ldots\ldots\ldots\ldots [1]$$

where e_1 = input, per cent
e_2 = output, per cent
E_{max} = saturation value of output, per cent
K = linear gain

The effect of amplifier-carrier action is included in the simulation shown in Fig. 3. Multiplier M_1 performs the modulation by obtaining the product of the input signal e_2 and the carrier signal $\cos \omega_c t$. The modulated carrier then passes through the amplifier block $G_1(\omega)$ which may contain time constants. For this analysis $G_1(\omega)$ is assumed to be a constant. The modulated signal $e_2 G_1(\omega) \cos \omega_c t$ then enters multiplier M_2, which demodulates the signal and gives an output of $e_2 G_1(\omega) \cos^2 \omega_c t$. This pulsating signal approximates the torque produced by the demodulating action of the two-phase motor when the flux contribution of the control

[1] Development Engineer, Leeds & Northrup Company. Assoc. Mem. ASME.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-9.

[2] "Optimum Nonlinear Control," by R. Oldenburger, Trans. ASME, vol. 79, 1957, pp. 527–546.

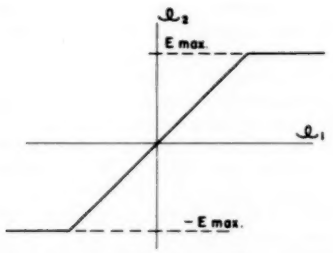


FIG. 2 AMPLIFIER SATURATION CHARACTERISTIC

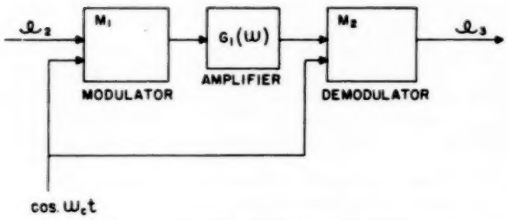


FIG. 3 CARRIER-AMPLIFIER SIMULATION

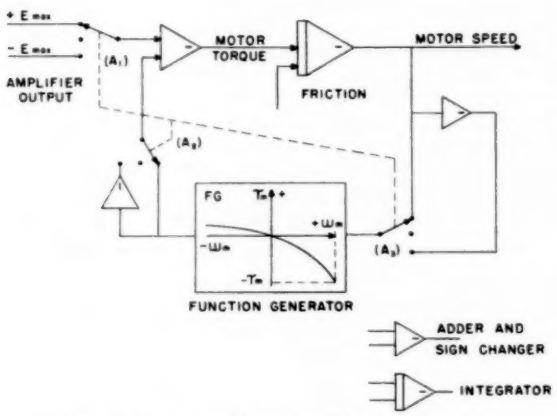


FIG. 4 SIMPLIFIED NONLINEAR MOTOR SIMULATION

winding is small compared to the flux contribution of the line winding. This condition is approached for small values of error signal and should be the most undesirable situation encountered.

Motor. A two-phase a-c servomotor is a nonlinear device which produces an output torque that is a function of output voltage and motor speed. The first part of this analysis assumes on-off operation of the motor and permits a simplified simulation using the actual speed-torque curve of the motor. Fig. 4 shows a block diagram of this motor simulation. The input signal to the motor is switched by relay contact A_1 while relay contacts A_2 and A_3 reflect the speed-torque curve for operation in the third and fourth quadrants.

When continuously variable motor operation is required, the simulation is simplified by using the conventional straight-line approximation of the speed-torque curve. The addition of a multiplier to Fig. 4 will make it possible to simulate the nonlinearities for variable input signals but this is not done in this paper.

Other Parameters. The various parameters such as motor and load inertia, gear ratio, dry-friction load, damping, and the input filter were made adjustable and could be varied over a wide range.

RESULTS OF COMPUTER ANALYSIS

Transient Response-Step Change. As was stated earlier in this paper it is important that the system exhibit the proper transient response to a step change of the input signal. The output of the recorder must not only reach the final balance point in the smallest possible time but it should come to balance with no overshoot.

For a servo such as the one considered here we should be able to obtain the fastest response without overshoot by running the motor with full output until a predetermined point is reached, then applying full reverse voltage to stop the system at the balance point, and then removing all voltage from the motor.

In order to determine the minimum time to respond to a 100 per cent step change in recorder input and come within about ± 0.2 per cent of balance, the computer is connected to give the on-off type operation described here. This not only gives a response curve that does not depend on damping methods but makes it possible to use the exact speed-torque curve of the motor.

The time required for the servo to reach balance for large step changes will depend upon several factors such as gear ratio, motor inertia, load inertia, load friction, and motor speed-torque characteristics. In this analysis the motor characteristics are held constant while the gear ratio is varied for selected values of load inertia and friction load. In all cases noted, there is an optimum gear ratio which will give the minimum balance time. When the gear ratio is too high the servo will velocity-limit and require an excessive time to travel across the chart. When the gear ratio is too low the servo will torque-limit and not come up to full speed. Fig. 5 shows these various types of response.

The optimum balancing times for various conditions are shown in Table 1. It can be seen that as load inertia or friction load is reduced, it is possible to reduce the gear ratio and obtain faster balancing times.

When the optimum gear ratio has been determined it is neces-

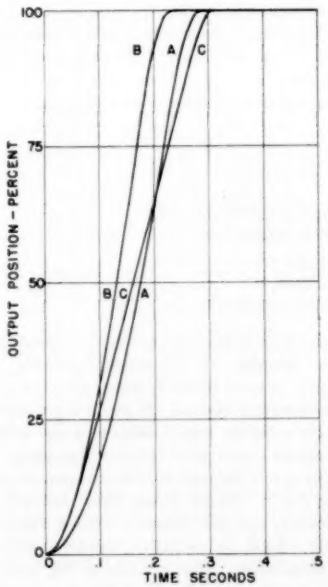


FIG. 5 TYPICAL SYSTEM RESPONSE TO 100 PER CENT INPUT-STEP CHANGE

(A, Gear ratio too low; B, correct gear ratio; C, gear ratio too high.)

TABLE 1 ONE HUNDRED PER CENT RESPONSE TIME WITH IDEAL DAMPING

Load inertia (oz-in.²)	19.2	19.2	9.60	9.60	9.60	6.40
Load friction (oz-in.)	15	10	15	10	0	10
Gear ratio						
4:1	...	0.290	...	0.222	0.185	0.200
5:1	...	0.255	0.240	0.203	0.180[a]	0.183[a]
6:1	0.270	0.240	0.220[a]	0.200[a]	0.180[a]	0.183[a]
7:1	...	...	...	...	...	0.190
8:1	0.255[a]	0.235[a]	0.220[a]	0.208	0.190	0.200
9:1	0.255[a]	...	...	0.216	...	...
10:1	0.263	0.245	0.240	0.227	...	...
12:1	...	0.265	...	...	...	...

[a] Minimum response time.

NOTE: Time in seconds, 100 per cent output change corresponds to 0.9 revolution (324°) of load shaft, inertia on motor shaft: 0.180 oz-in.²; servomotor: Diehl 5-watt, 2-pole, FPE 25-22.

sary to investigate how the system will perform for various damping networks and for various types of inputs. This requires a variable voltage to be applied to the motor rather than an on-off voltage. By using additional equipment it would be possible to include the effect of the nonlinear motor speed-torque curve, however, for the remainder of this analysis a linear speed-torque curve is used. The linear curve is chosen to give the same 100 per cent response as the nonlinear curve.

In order to investigate damping and other effects it is necessary to choose one set of conditions for gear ratio, inertia, and friction. For the remainder of this analysis the following conditions are assumed:

Gear ratio	8:1
Load inertia, oz-in.²	9.60
Load friction, oz-in.	10
Amplifier gain	full motor stall torque for 0.6 per cent error
Motor stall torque, oz-in.	5.5

The first means of damping to be considered is a passive lead-lag network. This has been a very popular damping method in the recorder field since the input filter also can be made to give this damping action.

This damping signal can be expressed by

$$e_d = \frac{\alpha T_s}{T_s + 1} c \qquad [2]$$

where

e_d = damping signal, per cent
c = system output, per cent
T = lag time, sec
αT = lead time, sec
s = Laplace transform operator

The performance of a lead-lag network depends upon α, the gain ratio of the network. If the value of α is less than 10, the network can give only a limited phase advance and the performance of the system is limited. When α is greater than 10, the network will give a better lead action, and the damping action obtained will almost equal pure velocity damping. Typical response curves for a 100 per cent step and values of α of 2, 4, and 10 are shown in Fig. 6. These curves show that it is not possible to obtain a critically damped response with α values of 2 and 4. These curves all exhibit an oscillatory response either giving an overshoot or undershoot. The α value of 10, however, gives a critically damped response. In all of these curves the value of damping time constant αT was adjusted to give the desired response.

For values of α of 10 or greater the results obtained with a passive lead network approach the results that can be obtained with velocity damping from a tachometer. With this type of damping there is no inherent time lag in the damping signal and a full 90-deg phase advance may be realized. Fig. 7 gives the response of the recorder with velocity damping for various sizes of input steps. In the first family of curves the damping-time constant is adjusted to give critically damped response for a 100 per cent step change. The second family of curves is for a system critically damped for a 10 per cent step change. The difference between the two systems can be easily seen. The system with optimum damping for a 100 per cent step is sluggish for smaller size steps, while the system optimized for a 10 per cent step has a pronounced overshoot for large step changes. This is a property of nonlinear systems with linear damping.

Although in most cases the response time for a 100 per cent step change is of primary importance and the system is adjusted to give critical response for this condition, it is interesting to investigate what results can be obtained with nonlinear damping methods. Oldenburger[3] describes a form of damping called "absquare damp-

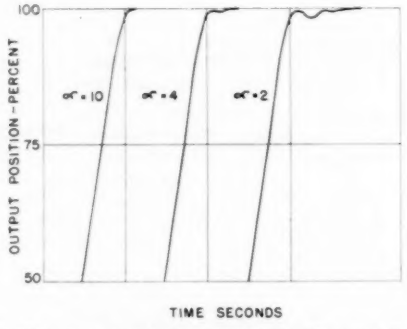


FIG. 6 PERFORMANCE OF LEAD-LAG-TYPE DAMPING CIRCUIT

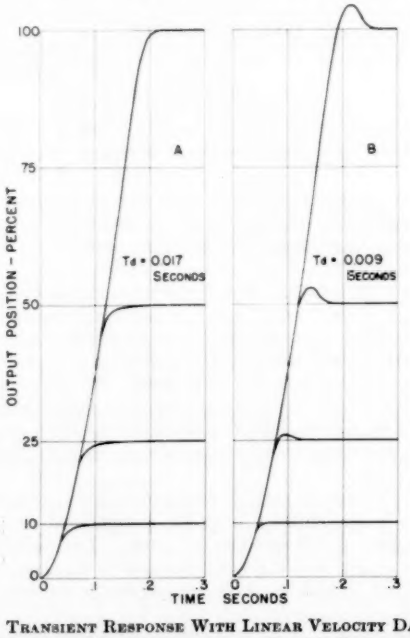


FIG. 7 TRANSIENT RESPONSE WITH LINEAR VELOCITY DAMPING (A, Critically damped for 100 per cent step; B, critically damped for 10 per cent step.)

ing" which uses a signal proportional to velocity plus a signal proportional to the velocity times the absolute value of the velocity.[3] This is given by

$$e_d = T_d \frac{dc}{dt} + \beta \left| \frac{dc}{dt} \right| \frac{dc}{dt} \quad \ldots\ldots\ldots\ldots [3]$$

where

e_d = damping voltage, per cent
T_d = linear damping time, sec
β = nonlinear damping time, sec²/per cent
$\frac{dc}{dt}$ = velocity of output, per cent/sec

With this type of damping the first term in Equation [3] provides the damping signal for small changes where the system acts approximately as a linear system. For larger signals where the system begins to approach velocity limiting, the second term in Equation [3] provides a large damping signal to reduce the overshoot present in Fig. 7, curve B.

Results with this type of damping are shown in Fig. 8. Here the system has a critically damped response for all sizes of input steps. The response time for small step inputs also is reduced.

In most practical industrial applications, it is necessary to incorporate an input filter in the system to reduce the effects of a-c pickup on the amplifier and motor. This filter must be designed to give the correct amount of attenuation to 60 cps a-c present in the input-signal source but the lag introduced by the filter must not affect the transient response excessively. Fig. 9 shows the response to a 100 per cent step of a typical input filter, the recorder system, and the filter and recorder together. If the time required for the input-filter response to reach within 0.2 per cent of balance is less than the balance time of the recorder, then it has no appreciable effect on the response to a 100 per cent step.

[3] "Combined Thyratron and Tachometer Speed Control of Small Motors," by A. J. Williams, Jr., Trans. AIEE, vol. 57, 1938, pp. 565–568.

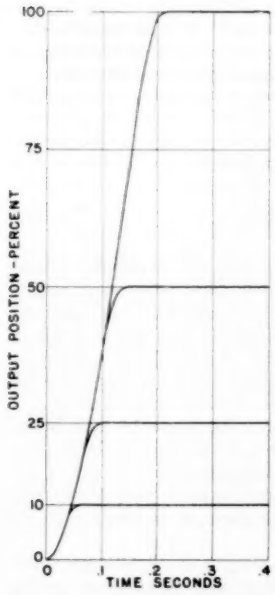


Fig. 8 Transient Response With Absquare Damping

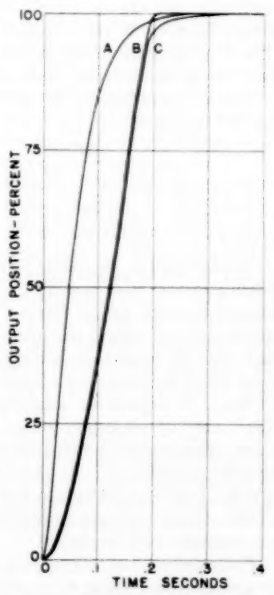


Fig. 9 Response of System With Input Filter
(A, Input filter only; B, recorder only; C, filter and recorder together.)

But if the response time of the input filter is greater than the response time of the recorder, then the recorder balancing time will be limited by the input filter.

It also should be noted that the input filter is a linear device with equal response times for large and small inputs. Thus even if the filter does not affect seriously the response time for large signals, it may tend to increase the response times for smaller changes and make the response times approximately equal for large and small signals. The lag introduced by the filter will then minimize any improvement from the nonlinear damping.

Transient Response Following Error. Although the response of this system to step changes is of primary importance, it is important to know how the system will respond to a constantly varying signal such as a ramp input. The recorder considered here is a Type 1 servomechanism[4] which exhibits a finite error when a constant-velocity input is applied. Fig. 10 shows typical responses for a constant-velocity or ramp input applied to the recorder. Here the output lags the input by a fixed time which is called the following error. This is also the reciprocal of K_v, the velocity constant of the system. For a linear system this value should remain constant regardless of the rate of change of the input signal. This also should hold true for the system considered here until saturation is reached. However, it will vary with the damping of the system. An overdamped system will have a greater following error than an underdamped system. Thus we should expect the recorder adjusted to give optimum response with a 100 per cent step change to have a larger following error than the one adjusted to give optimum response for a 10 per cent step change.

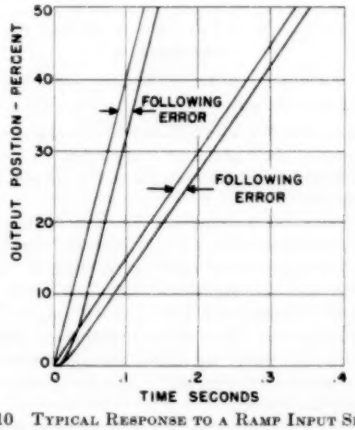


Fig. 10 Typical Response to a Ramp Input Signal

Table 2 Following Errors for Systems With Different Damping

Rate of input signal	Critically damped for 100% step	Critically damped for 10% step	Absquare damping
100%/sec	0.0175	0.01	0.0075
150%/sec	0.0175	0.01	0.0088
200%/sec	0.0175	0.01	0.010
400%/sec	0.0175	0.01	0.015

Note: All values in seconds.

Table 2 gives the values of following error for the systems with linear damping optimized for 100 and 10 per cent steps, and for the system with absquare damping. It shows that the following error is greater for the more highly damped system and that

[4] "Servomechanisms and Regulating System Design," by H. Chestnut and R. W. Mayer, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1951, pp. 194, 208.

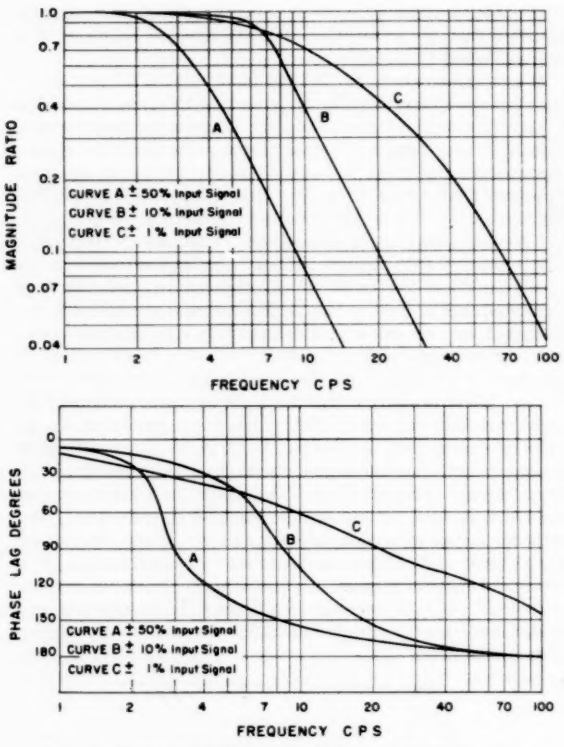


FIG. 11 FREQUENCY RESPONSE OF SIMULATED RECORDER

this error can be reduced when absquare damping is used. This would be expected since absquare damping gives a small linear damping signal for low velocities and a very large nonlinear signal for large velocities.

When an input filter is added to the system the following error introduced by the filter must also be considered. The total following error of the system will then be the sum of the recorder following error and the input-filter following error. In many cases the error introduced by the filter will greatly exceed the error from the recorder. Thus we see that if the requirements for a-c rejection are excessive, the lag of the input filter will overshadow the response of the recorder, and will greatly reduce the performance of the system.

When the effect of the carrier action is included in the simulation, no change in the transient response of the system can be seen.

Frequency Response. The response of this system to sinusoidal input signals is also important since many inputs can be approximated by a series of sine waves. A linear system would have one frequency response for all amplitude-input signals. However, this system, because of its nonlinearities, will have a family of frequency-response curves for various input amplitudes.

Fig. 11 shows the response of the simulated recorder for inputs of ±1, ±10, and ±50 per cent of full recorder range with the damping adjusted to give critically damped response to a 100 per cent step change. The curve for a ±1 per cent input signal closely approximates the response of a linear system. This would be expected since the amplitude is small and the system can operate over a semi-linear range.

The ±10 and ±50 per cent curves differ greatly from the linear curve. The magnitude ratio drops off more rapidly with a 2:1 slope and the phase lag increases sharply. This 2:1 fall-off of the magnitude ratio is caused by the saturation effect of the amplifier, which limits the voltage applied to the motor, thus limiting the torque available to accelerate and drive the load.

This saturation effect should not be confused with velocity limiting,[5] which has a frequency-response curve that drops off with a 1:1 slope. Velocity limiting occurs when a system can accelerate rapidly to full speed and then travel for some time. The simulated system, however, requires a rather long time to accelerate to full speed and consequently does not velocity-limit for most sine-wave inputs.

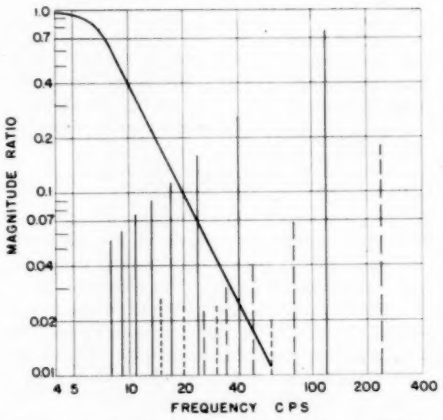


FIG. 12 FREQUENCY RESPONSE OF SIMULATED RECORDER WITH CARRIER EFFECT

The foregoing frequency-response analysis assumes no a-c carrier action in the amplifier. If this effect is included by using the simulation already described, the response curve shown in Fig. 12 is obtained. The frequency response for a ± 10 per cent signal with optimum damping for a 100 per cent step change is practically identical to the response obtained with no carrier action. The only exception is the presence of beat notes in the output of the system as the input signal near certain critical frequencies.

These critical frequencies for a 60-cps carrier can be described by the following

$$f = \frac{120}{n} \qquad n = 1, 3, 5, 7, \ldots .$$

$$f = \frac{240}{n} \qquad n = 1, 3, 5, 7, \ldots .$$

As the critical frequency is reached, the amplitude of these beats increases very rapidly to a maximum value which may be greater than the output signal. Fig. 12 also shows the maximum amplitude and critical frequency of the various beats. It can be seen that the magnitude of the principal group of beats, shown by solid lines, increases with frequency with a 1:1 slope and approximately equals the magnitude of the output signal at $f = 120/7$ cps. A secondary group of beats, shown by long dashed lines, whose critical frequencies are described by $f = 240/n$, have smaller magnitudes. At even submultiples of the carrier frequency, beat-type disturbances shown by short dashed lines occur in the output, but the magnitudes of these are relatively small and are constant with frequency.

If the presence of 60-cycle a-c pickup in the input signal makes it necessary to include an input filter with the recorder, the fre-

[5] "Dynamics of Electronic Self-Balancing Systems," by G. R. Jacob, paper presented at AIEE Conference on New Developments in Instrumentation, Boston, Mass., April 26–27, 1956.

quency response for small signals will generally be limited by the band width of the input filter.

Test Results

In an analog simulation it is desirable to have actual test results which can be used to check the results obtained from the computer. The operating parameters of the experimental recorder used to obtain this test data differ somewhat from the parameters used throughout this analysis. However, these differences will still allow comparisons to be made between the two systems.

Fig. 13 compares the transient responses for a 100 per cent step change of the actual recorder with no input filter and the computer simulation. The following operating conditions are used for the computer solution, since they more nearly approximate the conditions of the actual recorder:

Gear ratio	10:1
Load inertia, oz-in.^{2}	9.60
Load friction, oz-in.	15
Amplifier gain	full motor stall torque for 0.6 per cent error

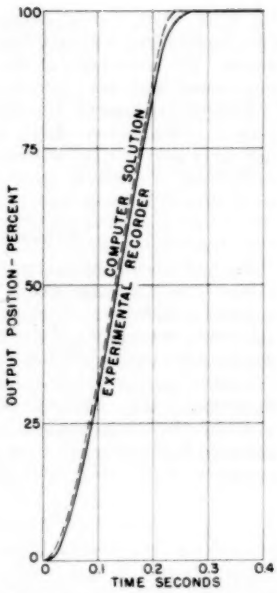


Fig. 13 Transient Response of Experimental Recorder and Computer Simulation

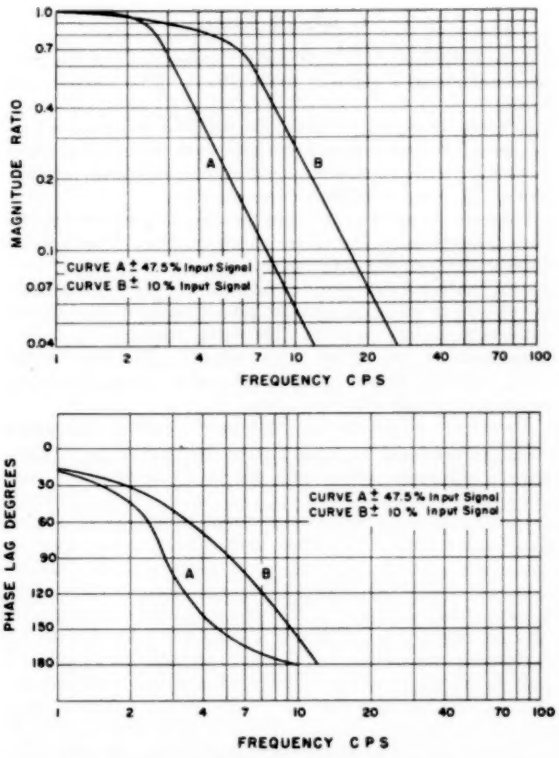


Fig. 14 Frequency Response of Experimental Recorder Without Input Filter

The response time of the actual recorder is slightly longer than the response time of the computer simulation and the final balancing of the recorder is more sluggish. This is primarily due to small differences in motor characteristics and damping networks between the computer simulation and the experimental recorder. The response times of the computer simulation and the recorder are 0.24 and 0.27 sec, respectively.

Fig. 14 shows the frequency response of the experimental recorder with no input filter for input signals of ±47.5 and ±10 per cent of recorder range. These curves are quite similar to the curves of the computer simulation shown in Fig. 11. The magnitude-ratio curve for the experimental recorder begins to drop off at a slightly lower frequency than the computer. This would be expected, however since the computer simulation has a full-scale balancing time that is slightly less than the full-scale balancing time of the experimental recorder. These curves also fall off with a 2:1 slope which is characteristic of a system with torque limiting.

The beat notes observed when carrier action is added to the computer simulation are also observed in the experimental recorder. However, beat notes are observed at many additional critical frequencies. The most prominent of these occurs at an input frequency of 30 cps and has a magnitude ratio of 0.3.

Conclusions

An analysis has been presented of a high-speed recording servomechanism. It has shown that for various combinations of load friction and inertia, there is an optimum gear ratio that will give a minimum response time for a 100 per cent step change in input signal. A value of gear ratio that is too high will cause the system to velocity-limit for long travels. A gear ratio that is too low will cause the system to torque limit.

Because of the nonlinear character of the problem, linear damping will not produce a critically damped response for all sizes of input step change. By using absquare damping, critically damped response can be obtained for all sizes of step inputs.

For large-amplitude sine-wave inputs, the system is torque limited and the magnitude ratio of the frequency-response curve decreases with a 2:1 slope which is characteristic of torque-limited systems.

If large amounts of 60-cycle pickup in the input signal make it necessary to use an input filter with too great a lag, both the transient and frequency response of the system will be severely limited.

When the effect of 60-cps carrier action is added to the simulation, no change is noted in the transient response. The frequency response is also unchanged except for the presence of beat frequencies in the output as the input frequency approaches certain critical frequencies. The experimental recorder exhibits beat notes at several additional critical frequencies, the most prominent of these being at an input frequency of 30 cps.

Although there are some differences between the two systems, the test results obtained from the experimental recorder are in good agreement with the results obtained from the computer simulation.

Discussion

A. J. Williams, Jr.[6] This paper clarifies many of the problems encountered in planning and building a high-speed recording device.

In Table 1 it seems that the number 0.225^a in the second column is in error. Perhaps it should be 0.255^a, as the number below it in the column, since the superscript "a" signifies that these two numbers are the ones for minimum response time.

Equation [3] gives the relation for damping, making use of a term proportional to the square of the velocity. This discusser likes to think of this term as useful because the motor has a limited braking torque and unlimited speed. The energy which the motor can absorb is therefore proportional to the remaining distance to the balance point. The energy stored in the motor is proportional to the square of the velocity. The latter energy which is kinetic should never be allowed to exceed the former energy if overshoot is to be avoided.

Two of the advantages for this type of damping are shown in the paper: First, shorter balancing times for the smaller step inputs, and second, smaller following errors for the smaller ramp inputs. The advantages of this type of damping become more conspicuous when even smaller steps and ramps are used.

A third advantage for this type of damping is indirect but important. Because less velocity feedback is used for small velocities, more gain can be used without inducing higher frequency oscillations. This greater gain reduces the position error which can result from dry-friction load.

[6] Science Director, Research and Development Dept., Leeds & Northrup Company, Philadelphia, Pa. Mem ASME.

Author's Closure

The author wishes to thank Mr. Williams for his interest in this paper. The observation about Table 1 is correct. A minimum response time for the second column of 0.255 sec is obtained for gear ratios 8:1 and 9:1. Mr. Williams' comments on absquare damping illustrate quite well how this form of damping action can act to prevent overshoot, regardless of the magnitude of the input step change. The third advantage of absquare damping given by Mr. Williams may be extremely important where high values of loop gain are required to minimize steady state position error.

Dynamic Study of an Experimental Pneumatic Process-Pressure Transmitter

By E. F. HOCHSCHILD,[1] PHILADELPHIA, PA.

A dynamic analysis and synthesis study of a highly accurate pressure-measuring instrument is presented. Emphasis is placed on methods, techniques, and the experience gained during the course of the investigation. Use of network theorems and the mobility method permits a simple analysis of a complex mechanical-pneumatic device. The effects on the transmitter resulting from tubing load are treated in detail. An electronic analog simulation, incorporating a transmitter nonlinearity, is used in the final synthesis for obtaining the required dynamic performance. Experimental and calculated frequency responses are compared and excellent correlation is shown.

Nomenclature

The following nomenclature is used in the paper:

A_b = rebalancing-bellows area, sq in.
A_i = input-diaphragm area, sq in.
$A_i^2R_i$ = input-damping coefficient, lb sec/in.
C_n = nozzle circuit volumetric capacitance, in³/psi
C_k = coupling capacitance due to pilot-valve motion, in³/psi
C_L = output-load capacitance, in³/psi
G_p = transfer function of pneumatic circuit
i = fluid flow through input capillary, in³/sec
K_1 = input-diaphragm grounded gradient, lb/in.
K_2 = series gradient, lb/in.
K_3 = grounded gradients other than K_1, lumped at input location, lb/in.
P_i = input pressure, psi
P_n = nozzle pressure, psi
P_o = output pressure, psi
R_1 = flapper-to-beam motion ratio
R_2 = feedback-force ratio
R_a,R_b,R_c = computer-circuit resistances, ohms
R_i = input-capillary resistance, psi/in³/sec
R_n = nozzle-circuit resistance, psi/in³/sec
R_p = pilot-valve resistance, psi/in³/sec
s = Laplace transform operator
T_1,T_2 = mechanical time constants, sec
x_1 = motion of input diaphragm, in.
x_2 = motion of primary beam at diaphragm location, in.
x_f = flapper motion, in.
Z_1 = impedance of input system, lb/in.
Z_L = driving-point impedance of transmitter load, psi/in³/sec
κ = open-loop gain
λ = flapper-nozzle gain, psi/in.
μ = pilot-valve gain

Introduction

Demands for improved control of industrial processes are often translated into more stringent requirements for measuring instruments. These requirements—accuracy, versatility, speed, and so on—in turn often impose difficult dynamic problems upon the designer. The solution of these problems may require the use of all available methods as well as the development of new techniques.

This paper presents a case study of the dynamics of an experimental process-pressure transmitter with a standard, 3–15 psig, pneumatic output. The primary purpose of the paper is not a detailed description of the work, but rather an exposition of the approach and method used in this study. Some fundamental questions, often only tacitly expressed during a dynamic investigation, will be answered in light of this specific study, in the hope of being of general applicability. The questions are: What is the relative value of analytical and test work? What information can be obtained by transient as compared to frequency-response tests? Of what value is an analog computer for this type of study? Is a really good correlation between test, analytical, and computer work possible?

The problem of obtaining good dynamic performance was primarily caused by high static-accuracy requirements and wide instrument rangeability. The transmitter, with an 8 to 1 span adjustability and a variable suppression, naturally had to perform well with any pneumatic-tubing load. The interrelation of these factors will be discussed in the paper. An additional restriction was imposed by requiring a minimum of difference between the pressure transmitter and a companion temperature-measurement instrument.

The dynamic analysis of the transmitter will be presented in two forms, one used for a computer simulation and the other, based on network theorems, used for graphical analysis. The major system nonlinearity will be included, and its modification of the otherwise linear analysis will be shown. Finally, results of the analysis and computer work will be compared to measured frequency and transient-response data.

Description of Transmitter

The transmitter, a pneumatic force-balance device, senses a process pressure and transmits a pneumatic signal proportional to the measured variable. Its output pressure depends upon the mechanical span and suppression adjustments. In operation, Fig. 1, the process pressure P_i acts on a seal diaphragm which serves as a barrier against process media. The signal is transmitted by means of a filled capillary system to a sensing diaphragm whose effective area is A_i. The force developed here is applied to the primary beam through a ball-and-seat arrangement and a short connecting screw. The resulting beam motion, which is amplified by the flapper linkage, causes a change in nozzle pressure P_n. The output pressure P_o, developed by the pilot valve, rebalances the input force by means of the bellows and secondary beam.

The instrument description, however, is incomplete without including the effect of the connected output load, the pneumatic

[1] Engineer—Control, Missile and Ordnance Systems Department, General Electric Company; formerly, Research Engineer, Minneapolis-Honeywell Regulator Company, Philadelphia, Pa. Assoc. Mem. ASME.

Presented at the Instruments and Regulators Division Conference, Evanston, Ill., April 8–10, 1957, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 4, 1957. Paper No. 57—IRD-7.

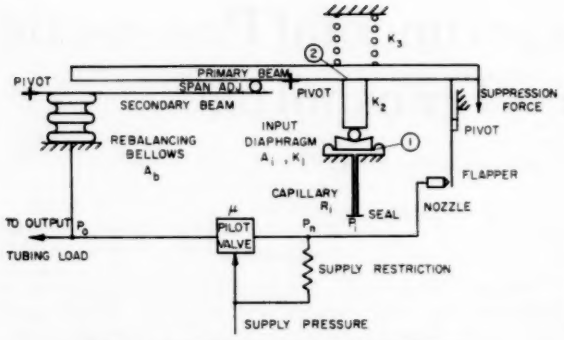


FIG. 1 SCHEMATIC DIAGRAM OF TRANSMITTER

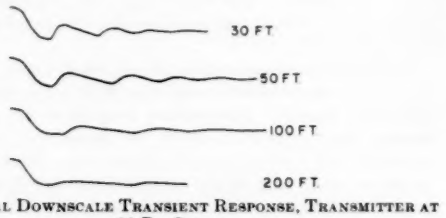


FIG. 2 INITIAL DOWNSCALE TRANSIENT RESPONSE, TRANSMITTER AT 20 PSI SPAN

transmission tubing. Its influence upon the dynamic performance of the instrument was clearly illustrated by early transient-response tests. In Fig. 2, the output response to downscale pressure-input steps, at the narrowest span, is sketched for several tubing lengths. Notice that for 200 ft the response was smooth, but for shorter lengths the response was oscillatory and somewhat nonlinear. The frequency-response data also exhibited a dependence on tubing load. However, the amplitude curve for medium tubing lengths showed no significant peak, an apparent contradiction to the transient data.

Comprehensive transient-response tests were made to scan the performance of the instrument for all operating conditions. These tests indicated the regions of worst performance, upon which future work had to be concentrated, and demonstrated that the dynamic response was affected by such variables as step direction, output-pressure level, and instrument span. The effect produced by increasing the span was especially difficult to understand since the response improved significantly at wider spans where the transmitter loop gain is higher!

Most of the effects described could have been suppressed by one of several cut-and-try methods. Each of these, however, would have slowed down the instrument. This was not deemed a satisfactory solution for it was desired to increase the speed of response of the transmitter. An analytical approach was therefore required, not only to explain the observed performance, but to find the best design solution.

LINEAR ANALYSIS

Using the instrument diagram, Fig. 1, as a starting point, a simple block diagram, Fig. 3, can be evolved. In the Appendix the transfer function of the filled capillary input system is derived. By taking a summation of forces at the diaphragm location, the motion x under the simultaneous action of the input pressure P_i and the feedback force F_o, referred to the diaphragm location, is given by

$$\frac{x}{A_iP_i - F_o} = \frac{1/K}{\frac{A_i^2R_i}{K}s + 1} \quad \ldots\ldots\ldots\ldots [1]$$

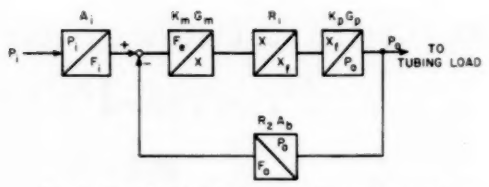


FIG. 3 SINGLE-LOOP ELEMENTAL BLOCK DIAGRAM

This first-order mechanical-component transfer function K_mG_m, where R_i is the capillary resistance and A_i^2/K the capacitance due to diaphragm motion, represents the damping effect of the filled input system upon the mechanical parts of the instrument. The inertia of the beams may be neglected; all spring gradients are referred to the diaphragm location and combined into one effective gradient. For simplification, the following terms are defined: The input force $F_i = A_iP_i$ and the error force $F_e = F_i - F_o$.

The beam motion x is amplified to flapper motion at the nozzle x_f by the lever ratio R_1. The linearized transfer function of the pneumatic circuit, K_pG_p, derived in the Appendix, is

$$\frac{P_o}{x_f} = \frac{\lambda\mu}{\frac{R_p}{Z_L}R_n(C_n + C_K)s + R_nC_ns + \frac{R_p}{Z_L} + 1} \quad \ldots\ldots [2]$$

This expression is more complicated than that derived by Gould and Smith (1)[3] or Helm (2) since it includes the coupling capacitance due to pilot input-diaphragm motion and the effect of tubing-load impedance. The driving-point impedance of tubing, represented here by Z_L, is discussed in detail in a paper by Rohmann and Grogan (3). It was shown that even in the simplest case the impedance is a capacitance, $\frac{1}{Z_L} = \frac{1}{X_L} = C_Ls$, and that it can assume very complicated forms. Equation [2], which is at best a simple second-order relation, shows how the tubing impedance affects the dynamic relation of the pneumatic circuit within the device loop.

The output pressure is fed back to the summing point by means of the bellows and the adjustable span-lever ratio R_2. As first and second-order terms exist in the loop, instability or poor dynamic behavior is possible.

On the basis of the block diagram, it was felt that open-loop transient tests could be run to determine accurately the loop gain and time constants for various tubing loads. This was done and a closed-loop frequency-response calculation was made. The results correlated very poorly with test data, even for zero tubing load. A subsequent accurate calculation of the component gains and time constants, which agreed fairly well with the open-loop tests, did not improve the correlation!

Hence it was obvious that, despite the open-loop agreement, the simple analysis did not represent the true closed-loop behavior of the device. This dilemma was finally resolved after a considerable amount of experimental work. Fig. 1 shows the ball-and-stem connection between the input diaphragm and the primary beam. The connecting parts are made of hardened steel and initially seated with a sharp blow so that a small area of contact is made. It had been assumed that negligible deformation would take place upon the application of external force; in other words, that an infinite series gradient existed. The experimental work proved that though this gradient was extremely large, it was not "infinite" and therefore could not be neglected. While this

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

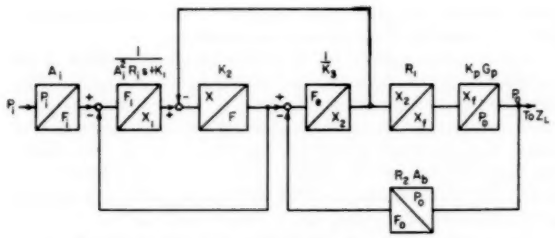


FIG. 4 MULTILOOP ELEMENTAL BLOCK DIAGRAM

factor resolved the puzzling problem, it complicated the analysis.

As shown in Fig. 1, the gradient of the diaphragm is designated by K_1, while K_2 is the rest of the grounded gradients[3] referred to the diaphragm location. The deflectional property of the stem is the series gradient K_3. The diaphragm motion x_1 resulting from the simultaneous application of the input pressure and stem force is an analogous expression to Equation [1].

Because of the existence of the series gradient, the block diagram now assumes the multiloop structure of Fig. 4. The difference in deflection, $x_1 - x_2$, between the lower end, 1, and the upper end, 2, of the stem causes a force F to be developed in the stem. This force is applied to the primary beam as well as to the input diaphragm. The motion x_2 of the primary beam results in a flapper motion and hence a change in output pressure. This pressure is fed back, as before, by means of the rebalancing bellows and secondary beam to the force-summation point on the primary beam, where $F_s = F - F_o$.

CLOSED-LOOP TRANSFER FUNCTION

The closed-loop transfer function can be derived from the block diagram by the simple, but laborious procedure of "collapsing" the loops until one main loop remains. There is a simpler expedient, however, which not only gives a single loop diagram, thus making a calculated analysis easier, but also gives a better "feel" for the device operation. The bases for this method are the important superposition theorem and the mobility concept of vibratory-system analysis.

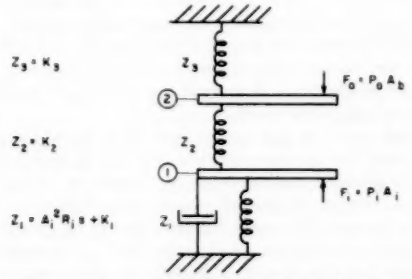


FIG. 5 IMPEDANCE REPRESENTATION OF MECHANICAL PORTION OF TRANSMITTER

The superposition theorem for a mechanical system may be stated as follows: "If a series of loads is applied to an elastic body and the deflection at any point in the body is considered, this deflection will be equal to the sum of the deflections at the same point that would be caused by applying each load individually to the body" (4). This theorem will be applied to the mechanical portion of the instrument, as represented in Fig. 5, in calculating the deflection x_2 under the simultaneous action of the input force F_i and the feedback force F_o. The mobility method (5) and its related impedance concept are useful aids for this calculation. In general, a mechanical impedance, that is, the force-displacement characteristics, is given by $Z = Ms^2 + Bs + K$, where M is the mass, B the damping coefficient, and K the spring gradient. For this system, the impedances are given by the simple expressions in Fig. 5.

To obtain the component of motion at 2 due to the action of the feedback force alone, the "self-impedance" Z_{22} (6) at this point must be calculated

$$Z_{22} = \frac{F_o}{x_{22}} = Z_3 + \frac{Z_1 Z_2}{Z_1 + Z_2} \quad \ldots\ldots [3]$$

To obtain the deflection at point 2 due to the input-force action at point 1, the resulting motion at point 1 is first calculated and then transferred to location 2. The force-displacement characteristic at point 1, the self-impedance Z_{11}, is given by

$$Z_{11} = \frac{F_i}{x_{11}} = Z_1 + \frac{Z_2 Z_3}{Z_2 + Z_3} \quad \ldots\ldots [4]$$

The transfer relation for motion is

$$x_{21} = x_{11}\left(\frac{Z_3}{Z_2 + Z_3}\right) \quad \ldots\ldots [5]$$

Substituting for x_{11} in Equation [4] gives the desired "transfer impedance" Z_{21}, for the motion at point 2 due to the input force

$$Z_{21} = \frac{F_i}{x_{21}} = Z_1 + Z_3 + \frac{Z_1 Z_3}{Z_2} \quad \ldots\ldots [6]$$

A new block diagram Fig. 6 can now be drawn, using the derived impedance relations Z_{22} and Z_{21}. Notice the simple structure which leads to a clear visualization and, hence, understanding of the device. To simplify the final derivation of the closed-loop transfer function, and clarify the instrument operation, the feedback loop can be reduced to unity. Fig. 7 shows the resulting block diagram.

The span of the transmitter is given by $A_i/(A_bR_2)$. The frequency-variant term, outside the loop, as derived in the Appendix, is

$$\frac{Z_{22}}{Z_{21}} = \frac{1}{T_2 s + 1} \quad \ldots\ldots [7]$$

This is a first-order lag whose time constant, $T_2 = (A_i^2 R_i)/K_2$, is governed only by the damping action of the input system and the magnitude of the series gradient. The impedance term

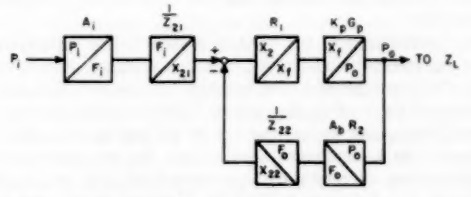


FIG. 6 SINGLE-LOOP IMPEDANCE BLOCK DIAGRAM

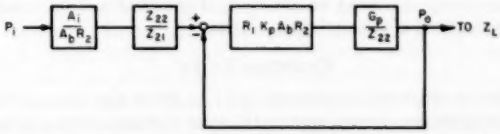


FIG. 7 REARRANGED SINGLE-LOOP IMPEDANCE BLOCK DIAGRAM

[3] A "grounded gradient" is a spring gradient referred to the instrument chassis or frame.

within the loop, also derived in the Appendix, is a compensated lag

$$\frac{1}{Z_{22}} = \frac{1}{K_1 + K_3}\left(\frac{T_2 s + 1}{T_1 s + 1}\right) \quad [8]$$

where $T_1 = (A_4^2 R_4)/(K_1 + K_3)$. The open-loop gain is then

$$\kappa = \frac{R_1 \lambda \mu A_5 R_2}{K_1 + K_3}$$

and the total open-loop transfer function is

$$(KG)_{OL} = \kappa G_p \left(\frac{T_2 s + 1}{T_1 s + 1}\right) \quad [9]$$

The closed-loop transfer function with unity feedback, as in Fig. 7, is the well-known

$$G_{CL} = \frac{(KG)_{OL}}{1 + (KG)_{OL}}$$

Substituting Equation [9], rearranging terms and then substituting Equation [2] for G_p

$$G_{CL} = \frac{1}{\frac{1}{\kappa}\left(\frac{T_1 s + 1}{T_2 s + 1}\right)\left[\frac{R_p}{Z_L} R_n(C_n + C_k)s + R_n C_n s + \frac{R_p}{Z_L} + 1\right] + 1} \quad [10]$$

The over-all instrument transfer function can now be obtained by combining the lag outside the loop with the closed-loop transfer function, G_{CL}. The final expression is

$$\frac{P_o}{P_i} = \frac{A_4/(A_5 R_2)}{T_2 s + 1} G_{CL} \quad [11]$$

The advantages of the foregoing method of analysis can now be seen quite clearly. Equation [11] shows that the dynamic response of the transmitter is dominated by a first-order lag which is independent of the pneumatic circuit, the grounded gradients, and the frequency-invariant gain terms. The closed-loop portion can easily be analyzed by standard graphical techniques applied to Equation [9] or by a direct numerical calculation of Equation [10], substituting $s = j\omega$. The effect of changes in gain or time constant upon device dynamics can be determined by a study of the open-loop frequency response. Even a complicated expression for the driving point impedance of tubing can be incorporated into the analysis in an exact, straightforword manner.

The disadvantage of the method is that device nonlinearities cannot be incorporated simply, nor can a study of parameter changes be accomplished very quickly. Because this was a developmental study of an instrument involving variations of tubing load, span, and so on, the use of an analog computer was indicated. As a significant nonlinearity, known to originate in the pilot valve, affected the output transients, Fig. 2, an analog simulation was definitely indicated. Component changes could then be made very quickly and dynamic-response improvement immediately evaluated on the basis of transient and frequency-response tests of the analog simulation.

Computer Study

For an effective computer study (7) a direct simulation of the transmitter was deemed essential in order to retain a close physical correspondence. It was therefore decided to use the "elemental" block diagram, Fig. 4, as a starting point instead of the diagram, Fig. 7, derived by the impedance approach.

A good representation for the dynamic characteristics of the pneumatic circuit, Equation [2] was of prime necessity. This presented several problems, the least of which were that pneumatic capacitance changes with pressure level and resistance with flow. The nonlinear pilot-valve flow resistance R_p, the related coupling capacitance C_k, plus the load-impedance effect were most difficult to simulate. This was solved by using a physical network for the pilot valve and its output load.

The computer diagram, utilizing conventional symbology (7), is shown in Fig. 8. Because a large-scale analog facility was not available at the time of this study, the circuit was obtained only after considerable effort. In order to reduce the number of amplifiers and obtain good scale factors the elemental block diagram was rearranged. By relocating area and gradient terms, all amplifier outputs were converted to pressure equivalents, and several amplifiers were eliminated.

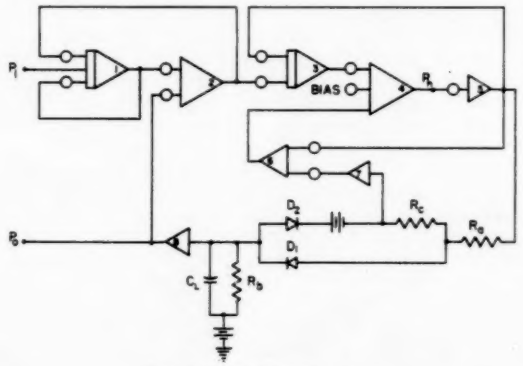


Fig. 8 Computer Diagram

As a result of the block-diagram manipulation all mechanical components, which formed the major part of Fig. 4, were reduced to amplifiers 1 and 2. The balance of the simulation is the pneumatic circuit. The physical network is an analog of the pilot resistance and output load. The coupling-capacitance effect, which depends directly on flow through the pilot valve, is represented by the voltage drop across the network resistors. "Unloading" amplifiers 7 and 8 prevent "nonphysical" current from being drawn from the network.

The pilot valve in this transmitter is a closed-loop device having a very low source resistance (R_a and R_e in Fig. 8), and therefore, high air-handling capacity. However, when flow is reversed, a change of the pilot-valve-stem forces must take place. To prevent the resulting dead spot in the pressure-flow relation from affecting the steady-state behavior a small, intentional bleed R_b is introduced. Diode D_1 in the network permits forward flow but blocks reverse flow. Diode D_2 acts in the opposite direction but will not conduct until the biasing voltage is overcome; this simulates the dead spot.

The tubing-load impedance Z_L is represented by a capacitor C_L, Fig. 8, as justified by reference (3) for the applicable frequency range. For short lengths of tubing it is completely valid. Even for a 50-ft tubing load the volume representation is good up to 100 cpm. This is shown in Fig. 9, reproduced from reference (3). For computer studies simulating 200-ft tubing load, the impedance was modified by including a series resistance.

With the exception of the pilot "dynamic" dead spot, the computer simulation was linear. To obtain parameter changes with pressure level or span, pot settings were altered for the particular

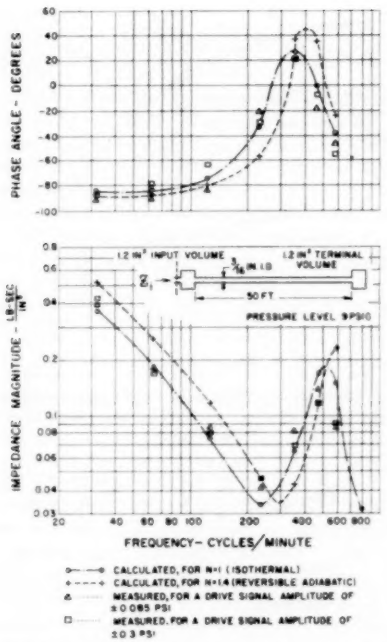


FIG. 9 DRIVING POINT IMPEDANCE OF 50-FT 3/16-IN-ID COPPER TUBING

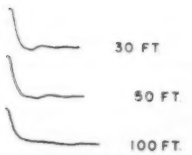


FIG. 10 FINAL DOWNSCALE TRANSIENT RESPONSE, TRANSMITTER AT 20 PSI SPAN

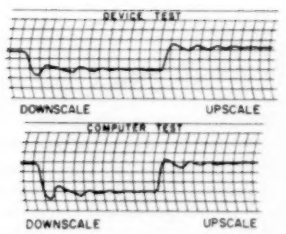


FIG. 11 TRANSIENT RESPONSE WITH VOLUME LOAD EQUIVALENT TO 50 FT OF 3/16-IN-ID TUBING, DEVICE AND COMPUTER TESTS

operating condition. By using mass-flow units certain simplifications were achieved. The computer scale factors were 100 v = 18 psi, 1 ma = 7.5 standard in³/sec, and 1 μfd = 0.6 in³.

The computer study was begun. After making slight "trimming" adjustments, excellent correlation with the device transient and frequency response was obtained for all tubing load. The purpose of the study then, was to find the changes which would most effectively produce faster dynamic response; minimize the variation due to tubing load; and suppress the pilot-valve nonlinearity.

It was immediately apparent that the two dominant loop lags, the mechanical lag and the nozzle time constant, must be "separated" to improve the oscillatory characteristic. Furthermore, the dominant mechanical lag had to be decreased to improve the instrument speed of response. When an apparently optimum response had been achieved, the changes were incorporated into the transmitter.

The transmitter, however, did not behave quite as expected. There were two main reasons for this. Smaller device lags, which previously had been negligible, had now become important, thereby causing a high-frequency oscillation. In addition, the volume representation of tubing load was no longer valid for the speeded-up device. Since there was no additional computer equipment available with which to improve the simulation, the final work was carried out experimentally on the transmitter.

By altering the pneumatic circuit and increasing the mechanical lag slightly, the high-frequency oscillations were removed and satisfactory transient response was obtained with all tubing loads. The resulting downscale transients, Fig. 10, can be compared to the initial tests in Fig. 2. Notice the faster response, the reduced load sensitivity, the removal of oscillations, and the almost complete suppression of the pilot-valve nonlinearity!

Although the device work was now essentially complete, one other task remained—to obtain a good calculated correlation for the frequency-response tests with tubing. Thus the validity of the analysis could be confirmed. It also would be possible to determine from this correlation whether the dynamic response of future similar devices could be predicted accurately before they were built.

Prior to making the calculation, numerical data were needed on the final device configuration. This could be obtained easily from the computer simulation provided the device response was matched. As the simulation of tubing load required more amplifiers than were available at the time, a simpler load had to be used. The transmitter was therefore tested with a volume-output load, even though this would not be a normal process installation. The computer circuit, with a capacitive load, was then adjusted until its transient response matched. As an excellent test comparison was obtained, Fig. 11, the simulated transmitter parameters could be used with assurance in the desired correlation study.

FREQUENCY-RESPONSE CORRELATION

For the purpose of the correlation study use was made of Equations [9], [10], and [11] which were derived by means of the impedance approach. The required numerical data were obtained from computer-potentiometer settings, tests of transmitter components, and open and closed-loop dynamic tests of the entire instrument. Consequently, several excellent cross checks were available.

First, the open-loop response, Equation [9], had to be calculated. The loop gain was established as $\kappa = 120$. The time constants of the compensated mechanical lag were found to be $T_1 = 3.3$ sec and $T_2 = 0.095$ sec. To correspond to the test data the transfer function of the pneumatic circuit, Equation [2], had to be evaluated for three tubing loads—dead-ended (representing 3 ft), 30 ft, and 200 ft. The dead-ended test data had been taken mainly to provide an experimental check for a calculated, linear transfer function. The dead-ended calculation was simple since Z_L was a pure capacitance and the pilot valve remained in its linear region of operation. G_p was found to be a second-order equation with two equal time constants of 0.04 sec.

For tubing loads of 30 and 200 ft the calculation of G_p turned out to be far from simple. Representing the pilot resistance R_p by some average constant value, and using a simple representation of tubing load resulted in large errors. First, to obtain good accuracy, the tubing-impedance expressions for the two lengths were derived from the theoretical transmission-line "propagation constants" in reference (3). Next, the effective pilot-valve resistance and its related coupling capacitance had to be determined as a function of frequency. And finally, the values of R_p, C_l, and Z_L for each frequency had to be combined with the constant R_nC_n to obtain the pneumatic-circuit transfer function, G_p.

The pilot-valve dynamic resistance had been determined by the describing-function method (8) in a previous analysis. The resistance, however, is not related to the output pressure but to

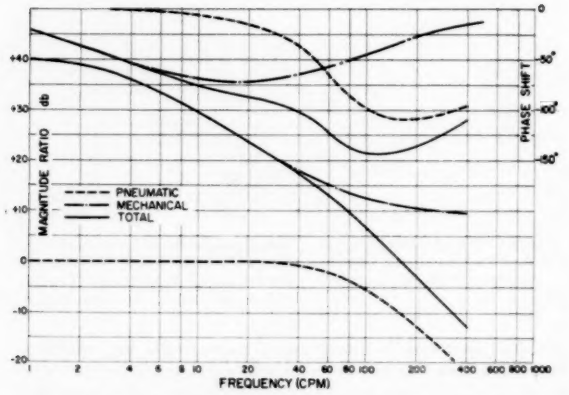


FIG. 12 OPEN-LOOP FREQUENCY RESPONSE OF TRANSMITTER WITH 30 FT OF TUBING LOAD

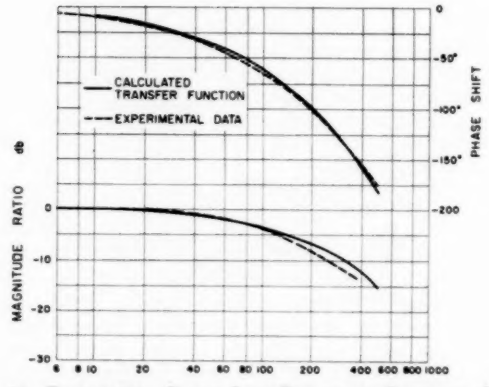


FIG. 13 TRANSMITTER CLOSED-LOOP FREQUENCY RESPONSE, CALCULATED AND EXPERIMENTAL, FOR DEAD-ENDED CONDITION (3-FT TUBING LOAD)

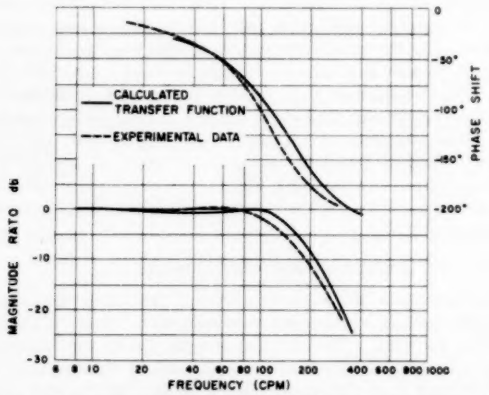


FIG. 14 TRANSMITTER CLOSED-LOOP FREQUENCY RESPONSE, CALCULATED AND EXPERIMENTAL, WITH 30-FT TUBING LOAD

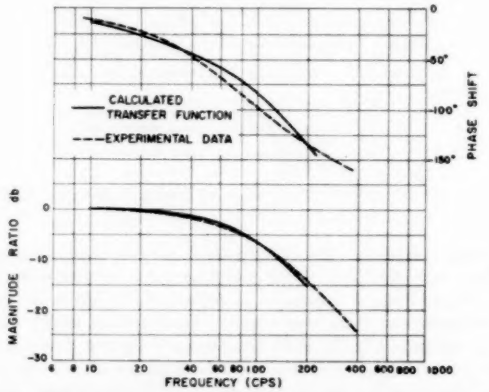


FIG. 15 TRANSMITTER CLOSED-LOOP FREQUENCY RESPONSE, CALCULATED AND EXPERIMENTAL, WITH 200-FT TUBING LOAD

output flow, which in turn can be calculated from $i_o = P_o/Z_L|_\omega$. Since P_0 is not known, *a priori*, this calculation would be one of successive trials involving the entire device transfer function. But there was a justifiable short cut available because test data of P_o as a function of frequency, had been gathered. From these data and the theoretical Z_L, i_o and in turn R_p, and C_k were calculated.

To obtain a satisfactory correlation the calculated pneumatic-transfer function had to be modified slightly by the inclusion of a small, effective load-separating resistance. This was justified for two reasons: (*a*) because additional resistance was introduced by fittings and valves present in the test setup; (*b*) but more important, because the tubing impedance was calculated from theoretical parameters, while the device test, with larger output pressure amplitudes, caused the tubing to behave in a nonideal, higher resistance manner.

The open-loop frequency response with 30-ft tubing load is plotted in Fig. 12. Notice that the phase curve of the compensated mechanical lag reaches a maximum and approaches zero at high frequencies. This effect, caused by the "separation" action of the series gradient K_2, is a stabilizing influence and explains why the initial calculation based on Fig. 3 was so much in error. Notice also the peculiar shape of the phase curve of pneumatic lag caused by the complicated dynamic relation for the tubing impedance. The total phase curve, which closely resembles a "conditionally stable" (9) dynamic system indicates

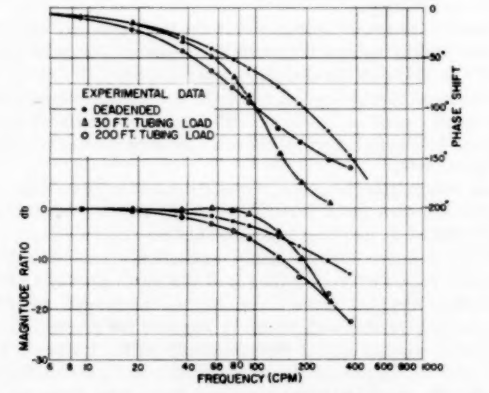


FIG. 16 FINAL FREQUENCY-RESPONSE TEST WITH TUBING LOAD, TRANSMITTER AT 20 PSI SPAN

why the initial, oscillatory transients, Fig. 2, for narrow spans were eliminated at wider spans with higher-loop gains.

The closed-loop response, including the effect of the load-separating resistor, was next calculated with the aid of a Nichols chart (10). The final transmitter frequency response was obtained by combining the closed-loop response with the mechanical

lag outside the loop. For ease of comparison the over-all calculated responses, as well as the test data for the three tubing loads, are shown in Figs. 13, 14, and 15.

The excellent correlation between the calculation and the test data is readily apparent for all tubing loads. This proves the validity of the analysis for a wide range of operating conditions. It is accurate enough to be used in the future for similar, theoretical calculations.

The effect of tubing load on the transmitter response is minimized by this design. This is readily apparent from Fig. 16, which shows the comparative response for the three widely different tubing loads. The smooth, gently sloping amplitude and phase characteristics would present no dynamic problem in the application of this transmitter to the control of a pressure process.

Conclusions

This paper has presented the methods used and the experience gained in a comprehensive dynamic investigation. It was found that the most important factor in reaching a successful conclusion is an accurate and detailed knowledge of the system under study. This applies not only to the over-all function and operation, but also to the static and dynamic characteristics of each and every component. Nothing should be assumed negligible, as was the series gradient in the connecting stem, until proved both experimentally and analytically.

The adequate corner frequency "separation" of adjacent dynamic lags, such as the mechanical lag T_1 and the nozzle lag, is necessary in order to obtain a well-damped dynamic response. Abrupt nonlinearities must be recognized and either eliminated or suppressed by suitable design. The range of operating conditions must be known or predicted so that the effect of these parametric changes can be considered in the analysis.

The paper has demonstrated the usefulness of applying the mobility method and network theorems to the analysis of a complex mechanical-pneumatic device. A considerably simpler analysis was achieved and a better understanding of the device operation was gained through the use of these tools.

It is felt that the answers to the four questions posed in the introduction have been answered successfully. To reiterate:

What is the relative value of analytical and test work? The paper has shown that they are interdependent and must be balanced carefully. Test work without analysis does not give a real insight into the mechanism of operation. On the other hand, analytical work without tests may mislead the investigator, either because too much may have been assumed or because numerical values may be significantly in error. Furthermore, to obtain a good calculated correlation, all static and dynamic measurements should be made on one device or system.

What information can be obtained by transient as compared to frequency-response tests? Both are necessary. In frequency-response data oscillatory tendencies may not be apparent because of the dominance of other lags, Equation [11], whereas these oscillations will be quite clearly shown by the transient tests, Fig. 2. Furthermore, as in this investigation, nonlinearities may be more evident in the transient tests. These, however, provide little information in the initial part, equivalent to the high-frequency spectrum, which is necessary for determining the order of the system. Finally, the laboriousness of calculating from transient to frequency response, and vice versa, can be avoided by testing for both.

Of what value is an analog computer for this type of study? The analog is a great time and laborsaving tool. Its usefulness, however, depends upon a careful analysis, and it does not replace thinking. After a correct analysis is established, a computer can save calculation time and speed up an investigation which is dependent on changing parameters. Furthermore, nonlinear effects which are always complicated and sometimes impossible to calculate, can be easily and completely represented on an analog computer. From an analog study a new approach for investigation often can be seen and then quickly explored. Finally, a computer can match the test results on a device or system and thereby provide an accurate knowledge of the numerical parameters.

Is a good correlation between test, analytical, and computer work possible? This paper proves that the answer is yes, providing that careful and correct work is done in all phases of the investigation. The comprehensive knowledge gained from such a study can be applied then to future extensions and investigations with full confidence.

Acknowledgments

The author appreciates highly the valuable assistance given to him, during the course of this study, by many members of the Research and Development Department of the Brown Instrument Division of Minneapolis-Honeywell Regulator Company; especially, C. P. Rohmann and E. C. Grogan for providing a significant portion of the background material and K. H. Stokes for his many contributions to the device phase of this investigation.

BIBLIOGRAPHY

1 "Dynamic Behavior of Pneumatic Devices," by L. A. Gould and P. E. Smith, Jr., ISA Conference Paper 52—9-2, Cleveland, Ohio, September, 1952.
2 "The Frequency-Response Approach to the Design of a Mechanical Servo," by H. A. Helm, Trans. ASME, vol. 76, 1954, pp. 1195–1214.
3 "On the Dynamics of Pneumatic Transmission Lines," by C. P. Rohmann and E. C. Grogan, ASME Paper No. 56—SA-1, unpublished.
4 "Elasticity in Engineering," by E. E. Sechler, John Wiley & Sons, Inc., New York, N. Y., 1952, p. 94.
5 "Mechanics of Vibration," by H. M. Hansen and P. F. Chenea, John Wiley & Sons, Inc., New York, N. Y., 1952, Chapter 6.
6 "Communication Engineering," by W. L. Everitt, McGraw-Hill Book Company, Inc., New York, N. Y., 1937, Chapter 7.
7 "Electronic Analogue Computers," by G. A. Korn and T. H. Korn, McGraw-Hill Book Company, Inc., New York, N. Y., 1952.
8 "Sinusoidal Analysis of Feedback Control Systems Containing Non-Linear Elements," by E. C. Johnson, Trans. AIEE, vol. 71, part II, July, 1952, pp. 169–182.
9 "Servomechanisms and Regulating System Design," by H. Chestnut and R. W. Mayer, John Wiley & Sons, Inc., New York, N. Y., vol. 1, 1951, p. 152.
10 Ibid., p. 319.

Appendix

Derivation of Transfer Function of Filled Capillary Input System

The expression for the motion of the input diaphragm as a result of the simultaneous application of the input pressure and the feedback force, referred to the diaphragm location, is derived in the following. For the capillary of the input system, Fig. 17, filled with an incompressible fluid, the laminar-flow pressure drop is

$$P_i - P_r = R_i i + L_i \frac{di}{dt} \quad \ldots\ldots\ldots\ldots [12]$$

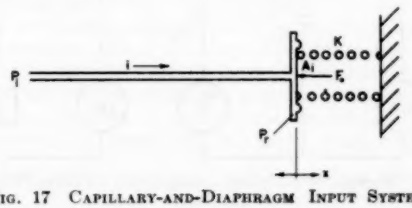


Fig. 17 Capillary-and-Diaphragm Input System

where R_i is the resistance of the tube and L_i is the inertance. The flow into the capsule which equals the flow through the tube, as the fluid is incompressible, is given by

$$i = A_i\,dx/dt \quad \ldots\ldots [13]$$

Differentiating both sides of Equation [13] yields the relation $di/dt = A_i\,d^2x/dt^2$. Substituting these two expressions into Equation [12]

$$P_i - P_r = R_iA_i\frac{dx}{dt} + L_iA_i\frac{d^2x}{dt^2}$$

Taking the Laplace transform, and setting the initial conditions equal to zero

$$P_i - P_r = (A_iR_is + A_iL_is^2)x \quad \ldots\ldots [14]$$

At the diaphragm location, the summation of forces can be written, where K is the effective grounded gradient referred to this point, $\Sigma F = P_rA_i - Kx - F_o = 0$

$$\therefore P_r = \frac{Kx - F_o}{A_i} \quad \ldots\ldots [15]$$

Substituting Equation [15] into Equation [14], multiplying by A_i, and rearranging

$$A_iP_i - F_o = (A_i^2L_is^2 + A_i^2R_is + K)x$$

For the case of interest, the inertance term was calculated to be negligible so that the final expression is

$$\frac{x}{A_iP_i - F_o} = \frac{1/K}{\frac{A_i^2R_i}{K}s + 1} \quad \ldots\ldots [1]$$

Pneumatic Circuit Linearized Transfer Function

The pneumatic-circuit transfer function relating output pressure to flapper motion can be derived from the linearized circuit, Fig. 18. The isolation amplifiers represent, respectively, the flapper nozzle and pilot-valve gain. In spite of the fact that the pilot is a closed-loop device, the added capacitance due to diaphragm motion, $C_k = A_p^2/K_p$, cannot be neglected in comparison to C_n. This complicates the analysis greatly. The circuit

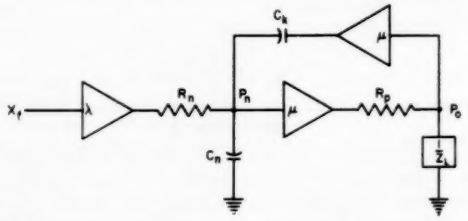


Fig. 18 Linearized Representation of Pneumatic Circuit

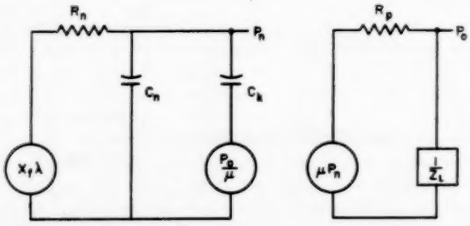


Fig. 19 Schematic Representations for Flapper-Nozzle Circuit and Pilot Valve

can be separated, Fig. 19, for the derivation of the transfer function, and the nodal equations can be written.

For the first circuit

$$(x_f\lambda - P_n)\frac{1}{R_n} = \frac{P_n}{X_n} + \frac{P_n - P_o/\mu}{X_k}$$

$$\therefore x_f\lambda = P_n[1 + R_nC_ns + R_nC_ks] - \frac{P_o}{\mu}R_nC_ks \quad \ldots\ldots [16]$$

For the second circuit

$$(\mu P_n - P_o)\frac{1}{R_p} = \frac{P_o}{Z_L}$$

$$\therefore P_n = \frac{P_o}{\mu}\left(\frac{R_p}{Z_L} + 1\right) \quad \ldots\ldots [17]$$

Substituting Equations [17] into [16] and rearranging, the final transfer function is obtained

$$\frac{P_o}{X_f} = \frac{1}{\frac{R_p}{Z_L}R_n(C_n + C_k)s + R_nC_ns + \frac{R_p}{Z_L} + 1} \quad \ldots\ldots [2]$$

Impedance Relations for Mechanical Components

Rewriting the impedance functions the following expressions are obtained

$$Z_{22} = Z_3 + \frac{Z_1Z_2}{Z_1 + Z_2} = \frac{Z_1Z_2 + Z_1Z_3 + Z_2Z_3}{Z_1 + Z_2} \quad \ldots\ldots [3]$$

$$Z_{21} = Z_1 + Z_3 + \frac{Z_1Z_3}{Z_2} = \frac{Z_1Z_2 + Z_1Z_3 + Z_2Z_3}{Z_2} \quad \ldots\ldots [6]$$

Then writing the term outside the loop and substituting the values for the impedances from Fig. 5

$$\frac{Z_{22}}{Z_{21}} = \frac{Z_2}{Z_1 + Z_2} = \frac{K_2}{A_i^2R_is + K_1 + K_2} = \frac{1}{\frac{A_i^2R_i}{K_2}s + \frac{K_1}{K_2} + 1} \quad \ldots\ldots [18]$$

As $K_1/K_2 \ll 1$, because the series gradient is very large, this term can be neglected, and substituting T_2

$$\frac{Z_{22}}{Z_{21}} = \frac{1}{T_2s + 1} \quad \ldots\ldots [7]$$

Substituting the impedance relations, the mechanical lag within the loop can be calculated

$$\left.\begin{aligned}\frac{1}{Z_{22}} &= \frac{Z_1 + Z_2}{Z_1Z_2 + Z_1Z_3 + Z_2Z_3} \\ &= \frac{A_i^2R_is + K_1 + K_2}{(A_i^2R_is + K_1)(K_2 + K_3) + K_2K_3} \\ &= \frac{\frac{A_i^2R_i}{K_2}s + \frac{K_1}{K_2} + 1}{(A_i^2R_is + K_1)\left(1 + \frac{K_3}{K_2}\right) + K_3}\end{aligned}\right\} \quad \ldots\ldots [19]$$

As also $K_3/K_2 \ll 1$, the expression can be simplified to its final form

$$\frac{1}{Z_{22}} = \frac{\frac{A_i^2R_i}{K_2}s + 1}{A_i^2R_is + K_1 + K_3} = \frac{1}{K_1 + K_3}\left(\frac{T_2s + 1}{T_1s + 1}\right) \quad \ldots\ldots [8]$$

The Time and Temperature Dependence of Thermal Stresses in Cylindrical Reactor Fuel Elements

By K. R. MERCKX,[1] RICHLAND, WASH.

A method of calculating the thermal stresses in cylindrical shapes is developed in this paper which uses a material model relating strain rate, temperature, strain, and stress. The material model, evaluated for unirradiated uranium, is used with this method to obtain the build-up and decay of the thermal stresses and strains in a solid-uranium fuel element operating in the temperature ranges of 100 to 350 C and 350 to 600 C during a 25-min period of increasing power generation and 700-min period of steady-state operation. During the period of increasing power generation, the elastic surface stresses are predicted to relax 60 per cent for the 100–350 C example and 48 per cent for the 350–600 C example. Further relaxation of 11.5 per cent for the 100–350 C case and 40 per cent for the 350–600 C case is calculated during the period of steady-state power generation.

Nomenclature

The following nomenclature is used in the paper:

E = elastic modulus, psi/per cent
σ = tensile stress, psi
ϵ = tensile strain, per cent
$C, m, c, H/R$ = parameters in material model
θ = temperature modified time, min
T = temperature, deg K
t = time, min
r, ϕ, z = polar radial, angular, and axial dimensions
G = elastic shear modulus, psi/per cent
α = coefficient of linear thermal expansion, per cent/K
$( )_p$ = plastic component
$( )_e$ = elastic component
$(\dot{})$ = differentiation with respect to θ
σ_i = principal stress component ($i = r, \phi, z$)
$e_T = \alpha \Delta T$ = strain due to thermal expansion

$$S^2 = \frac{1}{2} \sum_i s_i^2 = \text{second stress deviator invariant}$$

$$E_p^2 = \frac{1}{2} \sum_i e_{i_p}^2 = \text{second plastic-strain invariant}$$

$$I^2 = \frac{1}{2} \sum_i \left(\frac{de_{i_p}}{d\theta} \right)^2 = \text{plastic strain-rate invariant}$$

$$s_i = \sigma_i - \frac{1}{3} \sum_i \sigma_i = \text{stress-deviation components}$$

e_i = strain-deviation components

$$s = \frac{\sigma_r - \sigma_\phi}{2} = \text{maximum shearing stress in } r - \phi \text{ plane}$$

$$\gamma = \frac{e_r - e_\phi}{2} = \text{maximum shearing strain in } r - \phi \text{ plane}$$

$u = r/b$ = dimensionless radial distance
b = outer radius

Introduction

Reactor fuel elements generate heat by fissioning a small portion of the fuel material. In the operation of heterogeneous reactors using solid fuel elements, the fuel elements are required to maintain their physical integrity in order to keep the coolant from being contaminated with radioactive fission products. If severe ruptures occur, flow channels around the fuel elements may be blocked and structural damage may occur within the reactor. Though such failures can be expensive, savings in reactor size and fuel-inventory costs can be made if the fuel elements operate at their maximum power output. Thus the understanding of the generation and relaxation of thermal stresses and their contribution to the failure of reactor fuel elements is a fundamental problem in the economic design of a reactor.

Many of the possible fuel materials, such as uranium, are ductile metals. Hence any investigation of the mechanical failure of high-power-level fuel elements must consider the effects of plastic deformation, temperature, and stress relaxation on the thermal stresses and strains. The solution presented in this paper uses an approximation (1)[2] for the mechanical behavior of unirradiated uranium which relates the strain rate to the stresses, strains, and temperature. Such a material model can be used to show the effects of operating temperatures and the mechanical properties of the fuel material on the thermal stresses and strains.

The numerical results presented in this paper are for solid cylindrical fuel elements. Operating temperatures between 100 to 350 C and 350 to 600 C were assumed in order to demonstrate the effects of operation at different temperatures but similar power generation on the thermal stresses.

Method of Calculating Thermal Stresses

An accurate analysis of the thermal-stress condition in reactor fuel elements cannot be made until a model for the mechanical behavior of the fuel material is known. The radiation history of the fuel material, as well as the temperature, stress, and strain history will have to be factored into such a material model. Because of the experimental difficulties associated with obtaining the information required to evaluate the dependence of mechanical properties on irradiation, the model which follows is

[1] Research Engineer, Hanford Laboratories Operation, General Electric Company. Assoc. Mem. ASME.

Contributed by the Nuclear Engineering Division of The American Society of Mechanical Engineers and presented at the Nuclear Congress, Philadelphia, Pa., March 10–16, 1957.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 7, 1957.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

based on experimental data from unirradiated uranium. Thus the calculations presented in this paper approximate the stress conditions of a fuel element placed in a reactor for its first irradiation period.

The work of Dorn (2, 3) and Hollomon (4) suggested the form of functional dependence used in this study to relate the strain rate to the plastic portion of the strain, stress, and temperature. For a tensile specimen subjected to varying tensile stress σ, and temperature T, the strain at time t is given by the assumed mechanical model as:

Elastic Portion of the Strain

$$\epsilon_e = \sigma/E \quad [1]$$

Plastic Portion of the Strain

$$\frac{d\epsilon_p}{d\theta} = C\epsilon_p^m \sinh c\sigma \quad [2]$$

where the total strain ϵ is composed of an elastic and plastic portion

$$\epsilon = \epsilon_e + \epsilon_p$$

and the temperature and time are described by a temperature modified time

$$\theta = \int_0 \exp(-H/RT)dt \quad [3]$$

Though the preceding equations may not accurately predict the mechanical behavior for complicated loading histories, no better sets of analytical expressions were found at the time this analysis was started.

Creep tests conducted at Battelle Memorial Institute by F. R. Shober, L. L. Marsh, and G. K. Manning were analyzed (1) to obtain the following equations for the mechanical behavior of unirradiated uranium:

From 100 to 350 C

$$\frac{d\epsilon_p}{d\theta} = \exp(-30.03)\epsilon_p^{-11.5} \sinh(0.00148\sigma) \quad [4]$$

where

$$\theta = \int_0^t \exp(-13{,}500/T)dt$$
$$G = 7.3 \times 10^4 \text{ psi/per cent}$$
$$\alpha = 16.5 \times 10^{-4} \text{ per cent/K}$$

and from 350 to 600 C

$$\frac{d\epsilon_p}{d\theta} = \exp(63.7)\epsilon_p^{-0.587} \sinh(0.00188\sigma) \quad [5]$$

where

$$\theta = \int_0^t \exp(-60{,}600/T)dt$$
$$G = 2 \times 10^4 \text{ psi/per cent}$$
$$\alpha = 20.2 \times 10^{-4} \text{ per cent/K}$$
$$t = \text{min}$$
$$T = K$$
$$\sigma = \text{psi}$$
$$\epsilon_p = \text{per cent}$$
$$\alpha = \text{coefficient of linear thermal expansion}$$

In the work that follows, the material is assumed to be incompressible, homogeneous, and isotropic. Because the thermal stresses in cylindrical fuel elements are in a state of combined stress, Equations [4] and [5] should be expressed (5) in terms of the following stress and strain invariants

$$\sigma = \sqrt{3}\, S$$
$$\epsilon_p = \frac{2}{\sqrt{3}} E_p$$
$$\frac{d\epsilon_p}{d\theta} = \frac{2}{\sqrt{3}} I \quad [6]$$

where S, E_p, and I are defined in the nomenclature.

The equations of plastic flow (5) for the foregoing material model are

$$\bar{e}_i - \bar{e}_T = \frac{\bar{s}_i}{2G} + \frac{I}{S} s_i, \quad (i = r, \phi, z) \quad [7]$$

where the factor I/S is calculated using Equations [4], [5], and [6]. The term $(\bar{e}_i - \bar{e}_T)$ is the strain rate due to stresses, the $\bar{s}_i/2G$ gives the elastic portion of the strain rate, and Is_i/S gives the plastic portion of the strain rate.

The temperature and stress distributions are assumed to be independent of ϕ and z-axial symmetry and plane-strain conditions. The assumption of the plane strain is valid except near the ends of the fuel element if its length is several times greater than its radius. The σ_r, σ_ϕ, and σ_z are the principal stresses and ϵ_z is a function of time only. Using these restrictions, Equations [7] are reduced to

$$\bar{\gamma} = \frac{\bar{s}}{2G} + \frac{I}{S} s \quad [8]$$

$$\bar{e}_\phi = \frac{3}{2}\bar{e}_T - \frac{\bar{e}_z}{2} - \frac{\bar{s}}{2G} - \left(\frac{I}{S}\right) s \quad [9]$$

where

$$s = \frac{\sigma_r - \sigma_\phi}{2}; \quad \gamma = \frac{e_r - e_\phi}{2}$$

When Equations [8] and [9] are substituted into the compatibility equation

$$u\frac{\partial \bar{e}_\phi}{\partial u} - 2\bar{\gamma} = 0 \quad [10]$$

where

$u = r/b$ = dimensionless radial distance
b = outer radius of cylinder

the resulting differential equation for $\bar{s}$ is

$$\frac{1}{u}\frac{\partial(u^2\bar{s})}{\partial u} = 3Gu\frac{\partial \bar{e}_T}{\partial u} - \frac{2G}{u}\frac{\partial}{\partial u}\left(\frac{u^2 Is}{S}\right) \quad [11]$$

Equation [11] can be integrated and solved for $\bar{s}$. The result of this integration for a solid cylinder [$s(0) = 0$ is the boundary condition] is

$$\bar{s}(u) = \left(\frac{3G}{u^2}\int_0^u u^2 \frac{\partial \bar{e}_T}{\partial u} du\right) - 2G\frac{Is}{S} \quad [12]$$

The method of evaluating Equation [12] is suggested by the following interpretation: The first term on the right-hand side of Equation [12] gives the rate the stress $\bar{s}$ increases for an elastic material with the gradient of the thermal strain rate $\bar{e}_T$. The stress rate necessary to suppress the incompatible thermal expansions should be reduced if the material is allowed to relax plastically. The reduction of the stress rate by plastic relaxation is given by the term Is/S as would be expected for equations based on the theory of plastic flow; strain-rate terms are proportional to stress rate and a yield factor I/S times the stress. The equations of plastic flow are incremental in form; thus an incre-

mental type of solution, with the parameter θ determining the plastic-flow term, is used. When Equation [12] is written in incremental form, it becomes

$$\Delta_n s(u) = \frac{3G}{u^2}\int_0^u u^2 \frac{\partial}{\partial u}(\Delta_n e_T)du - 2G\left(\frac{Is}{S}\right)_{n-1}\Delta_n\theta \quad [13]$$

where

$$\Delta_n\theta = \int_{t(n-1)}^{tn} \exp(-H/RT)dt$$

$$\Delta_n e_T = e_{T(tn)} - e_{T[t(n-1)]}$$

and $(Is/S)_{n-1}$ is evaluated with the stresses and strains calculated from the previous steps.

Several of the additional terms which are needed to determine the quantity $(Is/S)_n$ are

$$\Delta_n\sigma_r(u) = 2\int_u^1 \frac{\Delta_n s}{u}\,du^*$$

$$\Delta_n e_s = 2\int_0^1 \left[\frac{\Delta_n s - \Delta_n\sigma_r}{3G} + \Delta_n e_T + \left(\frac{Is_z}{S}\right)_{n-1}\Delta_n\theta\right]udu^{**}$$

$$\Delta_n s_z = 2G\left[\Delta_n e_z - \Delta_n e_T - \left(\frac{Is_z}{S}\right)_{n-1}\Delta_n\theta\right]$$

$$S^2 = s^2 + \frac{3}{4}s_z^2$$

$$\Delta_n e_{i_p} = \left(\frac{Is_i}{S}\right)_{n-1}\Delta_n\theta$$

$$s_r = s - \frac{1}{2}s_z$$

$$s_\phi = -s - \frac{1}{2}s_z$$

$$(E_p)_n^2 = \frac{1}{2}\sum_{i=r,\phi,z}\left[\sum_{j=1}^{n}\Delta_j e_{i_p}\right]^2$$

For the initial increment

$$E_p = [(1-m)\Delta\theta C' \sinh cs]^{1/1-m}$$

and

$$e_{ip} = \frac{s_i}{S}E_p \qquad (i = r, \phi, z)$$

where the values of S, s_r, s_ϕ, and s_z are calculated by the foregoing method assuming $I/S = 0$. Further details describing this method of solution can be found in a document written by Merckx (6).

Numerical Calculations

The two cases selected to be calculated by the method of the previous section have temperature rises large enough to assure plastic yielding and equal temperature drops in two different temperature ranges so that the temperature-dependent effects of nearly equal power generation on the thermal stresses can be compared. The thermal stresses are calculated for solid cylindrical fuel elements assuming uniform heat generation, constant conductivity, and constant uranium surface temperature. With these assumptions, the temperature distribution for a solid cylinder is

$$T(u) = T(1) + \Delta T(1 - u^2)$$

The temperature drop is assumed to increase uniformly until some maximum where it remains constant—uniform increase of power generation until maximum power. For the purposes of calculation, the time dependence of T is assumed to be

$$\Delta T = 10t[\mathrm{K}],\ t \le 25\ \mathrm{min}$$

$$= 250[\mathrm{K}],\ t \ge 25\ \mathrm{min}$$

The cases of high and low-temperature operation are obtained by selecting the following uranium surface temperatures

$$\text{Low temperature} = T(1) = 373\ \mathrm{K}$$

$$\text{High temperature} = T(1) = 623\ \mathrm{K}$$

With the foregoing temperature distributions the increments of the thermal expansions become

$$\Delta_n e_T(u) = 10\alpha(1-u^2)(t_n - t_{n-1}) \qquad t < 25\ \mathrm{min}$$

$$= 0 \qquad t > 25\ \mathrm{min}$$

The foregoing values of the temperatures and thermal expansions are used with Equations [4], [5], [13], and the method described in the previous section to obtain the following values of the thermal stresses and strains: The calculated value of the maximum thermal equivalent tensile stress[3] for the cylindrical fuel element operating between 100 and 350 C is reduced from 90,200 to 36,400 psi by the stress relaxation during the period of power increase. After 700-min operation at steady-state thermal conditions, the surface stress is relaxed to a value of 31,200 psi or an additional reduction of 11.5 per cent. The maximum thermal equivalent tensile stress for the fuel element operating between 350 and 600 C is reduced from 30,200 psi to 15,600 psi during power increase and to 11,700 psi by stress relaxation after 700-min steady-state operating. The central stresses of the 350 to 600 C case, are completely relaxed out after 5 min operation at steady-state conditions. In Table 1 the values of thermal stresses calculated by the method described in this paper are compared to values of the thermal stresses calculated with elastic equations as well as with the theory of plastic deformations (7). The change of the stress distribution for the time-dependent solutions is shown in Fig. 1, while Fig. 2 shows the build-up and decay of the thermal stresses at the center and surface of the cylinder.

The plastic-strain distribution is also time and temperature dependent. Fig. 3 shows the dependence of plastic strain upon time and temperature for the two cases mentioned. In the low-temperature case, the maximum plastic strain occurs at the outer

Table 1 The Dependence of the Calculated Effective Stresses Upon the Assumed Material Model

		Case I	Case II
Surface temperature, deg C		100	350
Central temperature, deg C		350	600
Elastic solution (independent of time)	σ surface, psi	90200	30200
	σ center, psi	45100	15100
Plastic deformation theory	σ surface, psi	39400	22600
	σ center, psi	19700	11300
Time dependent (t = 25 min)	σ surface, psi	36400	15600
	σ center, psi	18400	920
With relaxation (t = 722 min)	σ surface, psi	31200	11700
	σ center, psi	14300	0

* The equilibrium condition for radial forces.

** The equilibrium condition of zero axial force.

[3] The equivalent tensile stress is that of the von Mises yield criterion for materials which are triaxially loaded and is equal to the tensile stress for simple uniaxial loadings. The equivalent tensile stress is $\sqrt{3}\,S$.

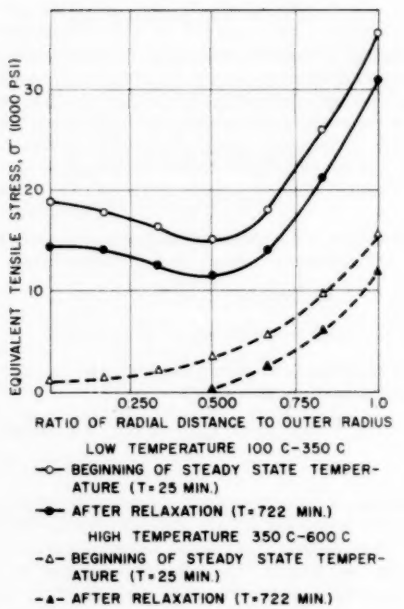


FIG. 1 EQUIVALENT TENSILE-STRESS DISTRIBUTION FOR A SOLID FUEL ELEMENT
(For two temperature ranges and different periods of stress relaxation.)

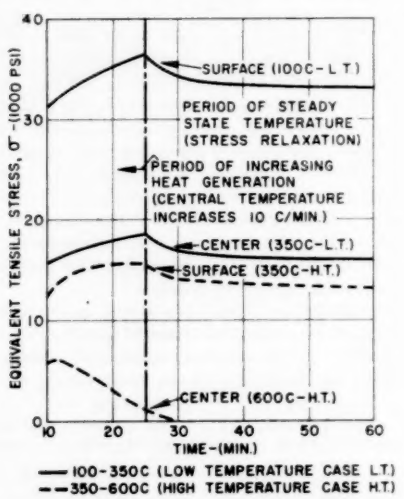


FIG. 2 EQUIVALENT TENSILE STRESS BUILD-UP AND DECAY FOR A SOLID FUEL ELEMENT
(Two temperature ranges.)

surface of the fuel element; while for the high-temperature case, the maximum plastic strain occurs in the center of the fuel element. In the latter instance, the outer surface strains are reduced by the increased straining of the hotter core material, and the increased central strains follow from the increased rate of stress relaxation in the center of the high-temperature cylinder.

DISCUSSION

The method presented in this paper can be adapted to stress calculations in any homogeneous, isotropic cylindrical object and

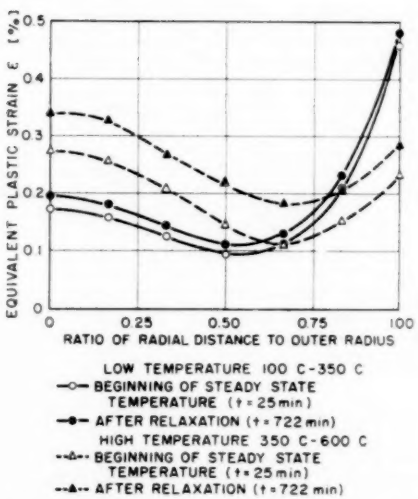


FIG. 3 EQUIVALENT PLASTIC-STRAIN DISTRIBUTION FOR A SOLID CYLINDRICAL FUEL ELEMENT
(For two temperature ranges and different periods of stress relaxation.)

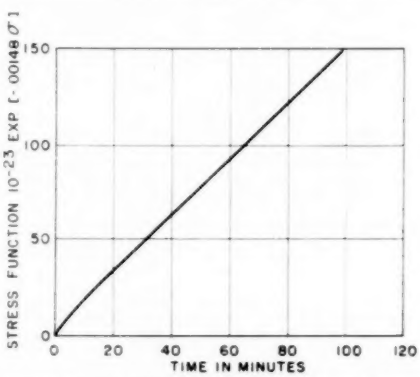


FIG. 4 STRESS RELAXATION DURING STEADY-STATE TEMPERATURES
(For surface stress of cylindrical fuel element operating between 100 and 350 C.)

could be generalized to any shape of body. Equation [11] which is for the thermal stress in cylindrical shapes can be evaluated for hollow cylinders by altering the constants of integration. Calculations for hollow cylinders also have been done with this method. Though the material model used in the numerical calculations is evaluated for uranium, it can be evaluated for any material if creep curves are available at two different temperatures. Adjustments of the parameters can be made to obtain the best fit for the experimental results. By altering the exponent m in Equation [2], the effects on the thermal stresses of different rates of strain-hardening of a material can be calculated. Changes also can be made in factor H in Equation [3] which will help predict the effects of having a different activation energy for the creep or stress-relaxation process. The factor c in Equation [2] determines the exponential dependence on stress, for high stresses the creep rate is exponentially dependent on stress. This exponential behavior on stress cannot be approximated by a viscoelastic solution and relate the observed stress dependence of creep curves; thus methods of analysis based on viscoelastic solution (8, 9) have to be modified if realistic metallic behavior is to be analyzed. In fact, the stresses will not relax exponentially as predicted by visco-

elastic solutions. For the 100 to 350 C material model, the surface stresses are relaxed according to the relation

$$\exp\left[-0.00148\sigma(t, 1)\right] = C't + \exp\left[-0.00148\sigma(0, 1)\right]$$

This equation assumes that there is little additional strain-hardening, which is true for the cylindrical fuel elements after constant power generation is reached. Fig. 4 is a graph of $\exp(-0.00148\sigma)$ versus t showing the validity of this assumption.

Because of the strong dependence of stress relaxation on the stress, numerical difficulties can be encountered in calculating the stress relaxation over a period of time using initial stress-relaxation rates and stress conditions. In certain cases of high stresses and high temperatures the relaxation is so rapid that the numerical calculations become unstable for increments of 0.1 min. In order to avoid these numerical difficulties, more complicated methods based on viscoelastic solution with time and temperature-dependent relaxation coefficients are being developed. These solutions will be incremental with the viscoelastic coefficients being altered with each increment.

Conclusions

Analysis of the numerical data obtained with the time and temperature-dependent model and supplementary calculations using elastic and plastic deformation theory gives the following general conclusions:

1 Assuming realistic temperature drops in massive uranium fuel elements, elastic calculations predict thermal-stress values which are several times too large.

2 For uranium fuel elements which operate in an intermediate-temperature range—100 to 400 C for unirradiated uranium—the initial thermal stresses and strains can be approximated with the theory of plastic deformations. The further relaxation of these thermal stresses, which may reduce these initial stresses up to 25 per cent, would have to be calculated with a rate-dependent-material model.

3 For high-temperature operation of fuel elements—over 400 C for unirradiated uranium—stress relaxation must be considered during power increase. For these cases, rate of power increase may affect the stress distribution. In fact, the thermal stresses may be largely relaxed during long periods of steady-state reactor operation.

4 During steady-state reactor operation, the stress dependence of the strain rate overshadows the effects of the changes in the strain rate due to the additional strain-hardening caused by the stress relaxation.

Hence once the initial stresses are calculated for the steady-state operating conditions, a material model considering only the stress and temperature dependence of strain rate could be used to predict the stress reductions due to stress relaxation.

Acknowledgments

The actual numerical evaluation of the resulting equations would be tedious if automatic computing techniques are not used. The author wishes to thank William C. McGee for programming this solution and his aid in solving the starting and stability problems which arose while making the initial calculations.

BIBLIOGRAPHY

1 "A Model of Mechanical Behavior Evaluated With Creep Tests Applied to Alpha Uranium," by K. R. Merckx, HW-40494, Office of Technical Services, November 17, 1955.

2 "Creep Correlations of Metals at Elevated Temperatures," by O. Sherby, R. Orr, and J. Dorn, *Journal of Metals*, vol. 6, January, 1954, pp. 71–80.

3 "What We Need to Know About Creep," by J. Dorn and L. Shepard, presented at the 57th ASTM Annual Meeting, June, 1954.

4 "The Flow of Metals at Elevated Temperatures," by J. H. Hollomon and J. D. Lubahn, *General Electric Review*, vol. 50, Parts I and II, February and April, 1947, pp. 28–32, 44–50.

5 "The Mathematical Theory of Plasticity," by R. Hill, Oxford University Press, London, England, 1950, Chapter 2.

6 "Thermal Stresses in Cylindrical Reactor Fuel Element," by K. R. Merckx, HW-42665, Office of Technical Services, June 4, 1956.

7 "A Variational Method for Determining the Thermal Stresses in an Infinite Cylinder," by K. R. Merckx, HW-31651, Office of Technical Services, March 24, 1954.

8 "Thermal Stresses in Thick-Walled Cylinders Exhibiting Temperature Dependent Viscoelastic Properties of the Kelvin Type," by H. H. Hilton, Proceedings of the Second U. S. National Congress of Applied Mechanics, 1954.

9 "Analytical Studies of Thermal Stresses in Media Possessing Temperature-Dependent Viscoelastic Properties," by H. H. Hilton, H. A. Hassan, and H. G. Russell, WADC-TR-53-322, September, 1953.